Arabinogalactan glycoprotein dynamics during the progamic phase in the tomato pistil

Authors: Cecilia Monserrat Lara-Mondragón and Cora A. MacAlister*

University of Michigan Department of Molecular, Cellular and Developmental Biology, Ann Arbor, MI, USA

* Corresponding author, macalist@umich.edu, ORCID: 0000-0003-1470-0596

Abstract

During fertilization, the male gametes are delivered by pollen tubes to receptive ovules, deeply embedded in the sporophytic tissues of the pistil. Arabinogalactan glycoproteins (AGPs) are a diverse family of highly glycosylated, secreted proteins which have been widely implicated in plant reproduction, particularly within the pistil. Though tomato (*Solanum lycopersicum*) is an important crop requiring successful fertilization for production, the molecular basis of this event remains understudied. Here we explore the spatiotemporal localization of AGPs in the mature tomato pistil before and after fertilization. Using histological techniques to detect AGP sugar moieties we found that accumulation of AGPs correlated with the maturation of the stigma and we identified an AGP subpopulation restricted to the micropyle that was no longer visible upon fertilization. To identify candidate pistil AGP genes, we used an RNA-sequencing approach to catalog gene expression in functionally distinct subsections of the mature tomato pistil (the stigma, apical and basal style and ovary) as well as pollen and pollen tubes. Of 161 predicted AGP and AGP-like proteins encoded in the tomato genome, we identified four genes with specifically enriched expression in reproductive tissues. We further validated expression of two of these, a Fasciclin-like AGP (*SlyFLA9*, *Solyc07g065540.1*) and a novel hybrid AGP (*SlyHAE*, *Solyc09g075580.1*). Using *in situ* hybridization, we also found *SlyFLA9* was expressed in the integuments of the ovule and the pericarp. Additionally, differential expression analyses of the pistil transcriptome revealed previously unreported genes with enriched expression in each subsection of the mature pistil, setting the foundation for future functional studies.

Key words: arabinogalactan proteins, tomato, pistil, pollen-pistil interaction, fertilization, glycoprotein

Key message: Pistil AGPs display dynamic localization patterns in response to fertilization in tomato. *SlyFLA9* (*Solyc07g065540.1)*, is a chimeric Fasciclin-like AGP with enriched expression in the ovary, suggesting a potential function during pollen pistil-interaction.

Introduction

Angiosperms constitute the vast majority of our economically important crops and the production of many of their seeds and fruits depends on sexual reproduction. Fertilization involves the delivery of the male gametes to receptive ovules, deeply embedded in the sporophytic tissues of the pistil. During this process, pollen tubes, carriers of the sperm cells, penetrate and elongate through the pistil tissues, establishing a number of interactions with distinct cell types along their journey (Palanivelu and Tsukamoto, 2012). To date, a number of ions and molecules including water, lipids, hormones, peptides and glycoproteins, derived from the pistil and controlling pollen germination and growth have been identified in different species (Cheung et al, 1995; Coimbra et al., 2007; Lush et al., 2000; Okuda et al., 2009; Pereira et al., 2016; Vogler et al., 2014; Wolters-Arts et al., 1998; Zheng et al., 2019).

Pollen behavior is modulated by the sporophytic tissues of the pistil, the stigma and style (pre-ovular guidance) and once the pollen tubes reach the ovary, they are further guided by signaling cues secreted by the female gametophyte or embryo sac (ovular guidance) (Higashiyama and Takeuchi, 2015). In the stigma and style, pre-ovular guidance is known to play an important role in maintaining interspecific barriers and, when the pollination is compatible, the stigma and style support pollen adhesion, hydration and germination, pollen tube elongation and directional growth, in addition to rendering pollen tubes competent to recognize ovular signals (Dresselhaus and Franklin-Tong, 2013; Kandasamy et al., 1994 ; Smith et al., 2013; Takeuchi and Higashiyama, 2011).

Reports across species have revealed that a family of highly glycosylated proteins, known as Arabinogalactan proteins (AGPs), play an important role during reproduction both in pre-ovular and ovular guidance. In the stigma, secretion of AGPs has been correlated to pistil maturity (apple – Losada and Herrero, 2012; magnolia – Losada et al., 2014; *Trithuria*- Costa et al., 2013; *Quercus*- Lopes et al., 2016) while in the style, members of this family such as the class III Pistil Extensin-like protein (PELPIII) and Transmitting Tract Specific (TTS) in tobacco promote pollen tube growth (Cheung et al., 1995; de Graaf et al., 2003,). In *Torenia fourneri*, an ovule derived arabinogalactan sugar (AMOR), likely derived from AGPs, is necessary to induce pollen tube competency to respond to ovule cues (Mizukami et al, 2016). In the ovary of *Arabidopsis*, a group of chimeric AGPs known as Early Nodulation-like proteins (ENODLs 11-15) are highly expressed in the embryo sac and the loss of function disrupts signaling necessary for bursting of the pollen tube upon penetration (Hou et al., 2016). Furthermore, an ovule expressed

classical AGP, AGP4/JAGGER is involved in promoting degeneration of the synergid cell post-fertilization and thus, blocking further pollen tube attraction (Pereira et al., 2016).

The structures and molecules involved in sexual reproduction tend to evolve rapidly (Edlund et al., 2004). This is likely true for the AGP family, as evidenced by the versatility of their functions in pollen guidance and the existence of lineage-specific members (e.g. PELPIII stylar proteins in *Nicotiana spp*.; Noyszewski et al., 2017). Therefore, in order to deepen our understanding of the extent of evolutionary conservation and/or emergence of novel AGP functions during sexual reproduction, it is necessary to broaden the spectrum of species studied. The study of species with relevance for agriculture, which, in many cases, remains understudied, is of particular importance due to increasing food demand and the yield sensitivity of crops to increasing global temperatures (Challinor et al., 2014). Tomato is an economically important crop and an emerging model system with structural features that other plant models lack (e.g. compound leaves, fleshy fruits, sympodial growth) (Kimura and Sinha, 2008). The tomato genome is publicly available (Sato et al., 2012) and the use of high throughput technologies to study its development, in addition to conventional cellular and genetics techniques, have shown to be extremely useful for identification of new molecular elements involved in floral meristem induction and fruit development/ripening (MacAlister et al., 2012; Pattison et al., 2016; Zhang et al., 2016). Regarding its reproductive biology, most efforts have been directed towards the study of fruit development and ripening, leaving a considerable gap of knowledge in our understanding of fertilization.

To date, despite the evidence pointing to a role for the AGP family in fertilization, their function in tomato reproductive biology prior to fruit development remains elusive. In this study, we explore the role of the AGP family during pollen-pistil interaction. First, using a series of probes for *in situ* AGP glycan detection, we characterize the spatiotemporal distribution of AGPs in pistils pre- and post-fertilization. Second, through transcriptome analyses, we evaluate the expression of predicted members of the AGP family in RNA-seq libraries derived from mature tomato pistils. Our search revealed a previously unidentified candidate of the AGP family with potential functions in pollen ovular guidance, setting the basis for future functional studies.

Materials and methods

Plant growth conditions

Tomato plants cv. Micro-Tom were maintained in a growth chamber under standard conditions (16 h light/18 h dark) until flowering. To avoid self-pollination in mature pistil samples, pre-anthesis flowers (7 mm length) were emasculated and allowed to further mature for 24 h. For pollen samples, the anthers of open flowers were dissected and placed in a microfuge tube with sterile dH_2O, and vortexed vigorously to release mature pollen, anther debris was removed with forceps, spun down and water removed.

In situ β-Yariv staining

To evaluate accumulation of AGPs in the stigmas of tomato pistils through development, two sets of five flowers in five different developmental stages (defined by flower bud length) were collected and sepals, petal and anthers removed under a dissecting microscope (Olympus, SZ61). The individual pistils were placed in a 0.2 mL tube upside down and the stigmas immersed in a solution containing 1mg/ml of β-Yariv or α-Yariv and incubated for 5 h at room temperature. The stigmas were then washed with 1x PBS and observed under the dissecting microscope to document accumulation of AGPs visualized as red-brown precipitates. For pistil cryosections, mature pistils were embedded in OCT medium (Tissue-Tek; Sakura Finetek USA) and flash frozen. Samples were sectioned 10 µm thick using a cryostat (Leica 3040S Cryostat) at -20°C. Sectioned samples were stained with 1mg/ml of β-Yariv or α-Yariv for 5 h and washed with 1x PBS. The samples were imaged using light microscopy (Leica, model DM5500B; camera Leica, model DFC365FX).

Pistil Aniline Blue staining

Unpollinated pistils and manually pollinated pistils collected after 24 h (five pistils per stage) were harvested and vacuum infiltrated with fixative (Acetic acid: EtOH, 1:3) for 24 h followed by 5 M NaOH overnight. The following day, the samples were washed for 10 min with dH_2O five times. Finally, the samples were incubated overnight with 0.001 mg/ml Aniline Blue Fluorochrome (Biosupplies Australia, 100-1) dissolved in 0.1 M K_2HPO_4 pH 10, mounted on an imaging chamber with Vectashield (Vector Laboratories, H-1400) and observed under an epifluorescence microscope (Leica, DM5500B) using the DAPI filter (355/455 nm) equipped with a Leica camera (DFC365FX).

AGP immunostaining

For immunolocalization of AGPs in semi-thin sections, we followed the protocol described by Costa et al., 2017. Briefly, unpollinated pistils and pistils 24 h after manual pollination (eight pistils per stage) were vacuum infiltrated with fixative (2% (w/v) Formaldehyde, 2.5% (w/v) Glutaraldehyde, 0.025 M PIPES buffer (pH 7.2) and 0.001% (v/v) Tween 20) for 2 h, the fixative was replaced with fresh solution and incubated overnight at 4°C. The samples were washed with 0.25M PIPES buffer pH 7.2 and dehydrated in ethanol series (25%, 35%, 50%, 70%, 80%, 90% and 3x 100%). After dehydration, sample embedding was performed by incubating the samples for 24 h in increasing concentrations of LR white (EMS, #14381) according to the manufacturer's recommendations. Following microtome sectioning to 200 nm, sections were incubated with blocking solution (filtered 5% non-fat skim milk in 1x PBS) for 10 min, washed with 1x PBS for 10 min and incubated with the primary antibody (JIM8, JIM13 or MAC207 1:5 hybridoma supernatant in blocking buffer) for 2 h at room temperature followed by an overnight incubation at 4°C. Sections were then washed with 1x PBS for 10 min and incubated with the secondary antibody (anti-rat-FITC, Sigma, F6258; 1:100 in blocking buffer) in the dark. The samples were then washed with 1x PBS, followed by a dH_20 wash. Calcofluor white was used as a counter stain. Slides were mounted with Vectashield (Vector Laboratories, H-1400) and observed under an epifluorescence microscope (Leica, DM5500B) equipped with a Leica camera (DFC365FX). Calcofluor white staining was visualized using the DAPI filter (355/455 nm) while FITC was observed using the 470/525 nm filter. False colored images were obtained with the Leica Application Suite (LASX) software.

For whole mount immunostaining of ovules, we followed the protocol described by Pasternak et al. (2015). Ovaries dissected from unpollinated pistils and pistils 24 h after manual pollination were fixed in vacuum with 2% formaldehyde in 1x Microtubule Stabilizing Buffer (MTSB; 50 mM PIPES, 5 mM $MgSO_4$, 5 mM EGTA pH 6.9) supplemented with 0.1% Triton x-100 pH 7, followed by a dH_20 wash. Samples were cleared by incubation with methanol at 60°C for 1 h. The concentration of methanol was gradually decreased by adding water to the samples every 5 min until the concentration of methanol was 20%, followed by a dH_20 wash. The samples were after incubated for 30 min at 37°C with 0.2% driselase and 0.15% macerozyme in 2 mM MES pH 5 to partially digest the cell walls of the samples and facilitate antibody binding. After cell wall digest, the samples were permeabilized with 3% IGEPAL and 10% DMSO in 1x MTSB. The samples were later incubated for 20 min with blocking buffer (2% BSA in 1x MSTB), followed by incubation with the primary antibody, MAC207 (1:5 in blocking buffer) for 2 h at 37°C. The samples were later washed three times with 1x MTSB and incubated with anti-rat-FITC (1:100 in

blocking buffer) for 2 h at 37°C. The samples were washed again three times with 1x MTSB, co-stained with Calcofluor white and finally mounted in an imaging chamber for confocal microscopy. Imaging was performed using a Leica Sp5 laser-scanning confocal microscope with 63x magnification. For FITC, the 488 nm excitation laser was used with a the RSP500 dichroic beam splitter and PMT detectors were set to capture light at a wavelength range of 495 to 554nm. For Calcofluor white, we used the 405 nm laser with the Substrat dichroic beam splitter, and PMT detectors were set at 414-474 nm wavelength. Z-slices were automatically optimized, and maximum intensity projections were generated using the LASX software.

mRNA sequencing

Mature pistils were fixed in ice cold acetone using a vacuum chamber followed by subdivision into four sections: stigma (n=35), apical style (1-1.6 mm below the distal end, n=35), basal style (3-3.5 mm from base, n=35) and ovary (n=35). For pollen samples, dry pollen was collected from mature open flowers and, immediately frozen with liquid nitrogen or incubated in liquid germination medium (PGM, recipe based on Covey et al. (2010), 24% w/v PEG 4000, 0.01% Boric acid, 2% Sucrose, 0.02% HEPES buffer, 0.003% $Ca(NO_3)\cdot 4H_2O$, 0.02% $MgSO_4\cdot 7H_2O$ and 0.01% KNO_3) for ~10 h, when the pollen tubes reached the tri-cellular stage (corroborated by DAPI staining in a small subsample). After incubation, the pollen tubes were spun down at 500 x g for 1 min, the PGM removed and the tubes immediately frozen with liquid nitrogen for RNA extraction. We performed a power analysis to determine the required number of biological replicates and sequencing depth to identify differentially expressed genes with at least a 2-fold difference and a p-value <0.01 (Busby et al., 2013) and data from tomato meristems (MacAlister et al., 2012) and pilot data from mature pollen grains. Four biological replicates of each sample were ground to a fine powder using a Tissue Lyser II (Qiagen, Cat. No. 85300). Total RNA was extracted using the RNeasy Plant kit (Qiagen, Cat. No. 74904) and treated in column with DNAse (Qiagen, Cat. No. 79254) according manufacturer's instructions. After total RNA isolation, the integrity of the samples was analyzed using a NanoDrop (Thermo Scientific) and TapeStation (Agilent). All samples except one stigma replicate passed quality control with A260/280 between 1.8-2.1, A260/230 >1.5 and RIN value ≥ 7.6, and were used for library preparation.

Stranded, poly-A enriched mRNA libraries were prepared by the University of Michigan Sequencing Core using the TruSeq mRNA stranded library prep kit (Illumina, Cat. No.

20020594). The samples were multiplexed and 50 bp, single-end sequencing was performed in the Illumina HiSeq4000 platform. We performed quality control of the de-multiplexed reads using FastQC (Wingett and Andrews, 2018). The reads were filtered based on quality and trimmed (Trimmomatic, Bolger et al., 2014) of any adapter contamination. Transcript abundance was quantified using kallisto (Bray et al., 2016) pseudoalignment, using the transcriptome of the tomato cv. Heinz 1706, ITAG 3.2 version to build the index (Sol Genomics Network; Sato et al., 2012; supplemental table 1).

Data mining and exploratory analyses

To complement our data, we compiled publicly available datasets from different sources including Ezura *et al*. (2017). The samples correspond to a variety of vegetative tissues collected from tomato cv. Micro-Tom plants, including cotyledon, young, mature and old leaves, 7 days old and mature root and stems. Additionally, datasets corresponding to dissected anthers in three different stages, sepal, petals, and green and red fruits. Additionally, we included in our analysis datasets of mature leaf (2 biological replicates), root (2 replicates) and flower bud (2 replicates) from the cv. Heinz; deposited in the SolGenomics Network (SGN) website (Sato et al., 2012). The raw reads from the aforementioned samples were downloaded from the NCBI database and processed as above. The data was filtered based on their TPM values with genes with TPM higher or equal to 0.1 considered as expressed. After filtering, the data was transformed by the Variance Stabilizing Transformation method (VST) using the *vst* function of the DEseq2 package (Love et al., 2014). The Principal Component Analysis (PCA) was carried out in R studio using the *plotPCA* function of DEseq2. To determine transcriptome complexity (Carnici et al., 2000), the average expression of each gene across biological replicates from the same tissue was calculated. The contribution of each gene to the overall transcriptional load of each tissue (Total Transcriptional Output, TTO) was calculated by sorting the data from highest to lowest expression value and dividing each value by the sum of all average expression values. The cumulative distribution of the transcriptional contribution of each gene and standard deviation was plotted.

Tissue exclusivity and preferential expression

Genes with preferential expression across reproductive tissues were determined using DEseq2 (Love et al., 2014) by performing pair-wise comparisons of the normalized data between each

tissue versus the rest of the tissues. For the differential expression analysis, we used the datasets from SGN with their respective biological replicates (leaf, root and flower bud). From the datasets of the study of Ezura et al. (2017), we only included the datasets from other floral whorls but pistil (sepal, petal and anther). Genes with False Discovery Rate (FDR) q-value < 0.05 and Log Fold Change (LFC) > 2 were considered as preferentially expressed. Next we looked for overlap of the resulting list of genes among tissues and retained the unique elements to each tissue/structure for further analyses. To predict the subcellular localization of the enriched genes, we used DeepLoc-01 (Almagro Armenteros et al., 2017) and the amino acid sequences of the enriched genes.

AGP expression, quantitative RT-PCR and in situ hybridization

ITAG identifiers of predicted AGP members (161 genes), including canonical AGPs, chimeric and hybrid AGPs were compiled from previously published reports (Ma et al., 2017; Johnson et al., 2017, Showalter et al., 2010). A matrix of the log2 of the mean TPM values per gene was plotted as a heatmap using the *pheatmap* R package. Genes were clustered based on their expression patterns through calculation of Euclidean distances. Out of the initial list, 60 AGP genes were represented in our transcriptomes and the published vegetative transcriptomes. We then selected genes with a high expression in pistil tissues compared to the rest of the tissues for independent validation by quantitative RT-PCR (qRT-PCR). Total RNA was isolated as above from mature leaves, sepals, petals, flower buds (2 mm length), anthers, pollen, ovaries, basal and apical style and stigmas. DNA-free total RNA (200 ng) was used for cDNA synthesis using the SuperScript® III First-Strand Synthesis System (Invitrogen, 18080051) following manufacturer's instructions. Serial tenfold dilutions of pooled cDNA from the different tissues were used to determine primer efficiency curves. Quantitative PCR was performed using the Applied Biosystems® SYBR® Green PCR Master Mix in a StepOnePlus™ Real-Time PCR System according to the manufacturer's manual. Three independent biological replicates were analyzed per sample per gene using the primers in supplemental table 2. Actin and ubiquitin were used as reference genes and the obtained data was normalized using the Livak calculation method (Livak and Schmittgen, 2001). For *in situ* hybridization of mature pistil sections, we followed the protocol published by Javelle et al. (2011) using sense or antisense DIG-labeled RNA probe generated from the full *SlyFLA9* cDNA sequence amplified with the primers in supplemental table 2.

Sequence homology search and phylogenetic analysis

The Basic Local Alignment Search Tool (BLAST) available in the SolGenomics Network (SGN; https://solgenomics.net/tools/blast/) website or The Arabidopsis Information Resource (TAIR) was used to identify homologous sequences (supplemental table 3). Multiple sequence alignment was performed using Clustal Omega. The phylogeny of Arabidopsis and tomato GALT/HPGTs was generated with Phylip using the neighbor joining method with 200 bootstrap replicates; the consensus tree was calculated by the extended majority rule.

Results

*Arabinogalactan glycoproteins accumulate in mature tomato pistil*s

Secretion of AGPs to the stigmatic surface correlates with pistil receptivity in several species with wet stigmas (Gell et al., 1986; Losada and Herrero, 2012; Losada et al., 2014; Losada et al., 2017; Lord and Heslop-Harrison, 1984). We asked whether this observation also holds true in tomato pistils, another species that possess a 'wet' stigma upon maturation (Heslop-Harrison and Shivanna, 1977). The stigmas of pistils from five different developmental stages (defined by the size of the flower bud) up to anthesis were submerged briefly in a solution containing the AGP-specific synthetic dye, β-Yariv (Yariv et al., 1962). As a negative control, we used the non AGP-reactive isoform α-Yariv in pistils of the same developmental stages. As shown in Fig. 1A, a progressive accumulation of the β-Yariv-AGP complex precipitates in the stigma as the pistil reaches maturity, while the negative control does not display any visible staining. Our observations support the hypothesis that pistil receptivity correlates with secretion of AGPs in the stigma of the tomato pistil. We then asked whether AGPs reactive to the β-Yariv reagent are also present in the style and ovary. The transmitting tract of the style is a specialized secretory tissue rich in carbohydrates, flavanols, glycoproteins and lipids (Lennon et al., 1998), while the ovary contains the radially arranged ovules which are attached to the placenta (Figs. 1B, 2). Frozen sections of mature pistils were stained with β-Yariv and α-Yariv and observed under the microscope. Compared to the negative control, clear β-Yariv staining was observed in the transmitting tract of the pistil and ovules, indicating that AGPs are produced by different cell types along the pistil (Fig. 1B).

To expand our observations with the β-Yariv reagent, we generated semi-thin sections of tomato pistils and probed them with anti-AGP glycan antibodies: JIM8 (unknown epitope

structure), JIM13 and MAC207 (described epitope for both antibodies: β-GlcA1 → 3-α-GalA1→2-Rha) (Knox et al., 1991; Pennell et al. 1991; Yates et al., 1996). Even though all of these antibodies bind to AGP glycan moieties, they display differential binding and developmental patterning (Knox, 1995). Several studies have also shown AGP localization to be responsive to various physiological and developmental events such as biotic and abiotic stress (Wu et al., 2017), and during male and female gametophytic development (Acosta-García and Vielle-Calzada, 2004; Coimbra et al., 2007). Particularly in tomato, reports of AGP redistribution during fruit ripening and softening (Leszczuk et al., 2020) and during fruit development under abiotic stress (Fragkostefanakis et al., 2012; Mareri et al., 2016), support the hypothesis of dynamic spatiotemporal regulation in response to stimulus in tomato. We, therefore, focused on detecting AGP localization changes during pollen-pistil interaction by collecting pistils before and after fertilization. To determine an appropriate post-fertilization time point, we used aniline blue staining to visualize the position of pollen tubes within the pistil tissues. We found that at 24 h post pollination, pollen tubes had entered the ovary and successfully reached several ovules (Sup. Fig. 1). In contrast to our observations with β-Yariv staining, unpollinated pistils displayed weak labeling with all three antibodies (Sup. Fig. 2), suggesting that these structures produce AGPs which are not recognized by, or with weak affinity to the tested antibodies, though it is also possible that the antibody epitopes may be present, but masked in these samples (Rydahl et al., 2018). Papillae in the stigma and cells along the transmitting tract are highly secretory and the complexity of their extracellular matrices might limit the accessibility of the antibodies to their epitopes. Weak immunolabeling with the antibodies in the stigma cells and style was also observed in samples at several stages during pollen-pistil interaction while strong labeling was observed in pollen grains attached to the papillae cells (Sup. Fig. 2C, MAC207 antibody). In ovary sections, on the other hand, unpollinated pistils displayed a punctate, intracellular labeling in the ovules when stained with JIM8 and JIM13 (Fig. 2A and B). In tomato pistils, due to the size of their ovules, obtaining sections of the ovary that preserve the embryo sac intact is challenging. We were able, however, to observe strong signal in a restricted area towards the micropylar pole of the ovule when staining with MAC207 (Fig. 2C); potentially labeling the synergid cells (Fig. 3C). After pollination, all three antibodies strongly labeled the cell walls of the inner cell layer of the integuments, surrounding the embryo sac (Fig. 2D to F).

To have a better visualization of the ovule surface and expand our characterization of the ovular AGP patterning, we performed whole mount immunolocalization with MAC207. Whole mount immunostaining was performed in ovules from the same time points (unpollinated and 24 hap)

as above. Our results show that, in ovules from unpollinated pistils, a subset of MAC207-reactive AGPs were restricted to the micropyle (Fig. 3A). Interestingly, the signal visible in the micropyle of unpollinated ovules was no longer visible at 24 h after pollination (Fig. 3B). AGP labeling by the MAC207 antibody was also present in the pollen tube (Fig. 3B).

Identification of genes with enriched expression in tomato reproductive tissues

In tomato, a number of studies have focused on identifying sets of genes with potential function during fruit development (Pattison et al., 2015; Ezura et al., 2017; Zhang et al., 2016); however, a significant gap of knowledge still exists in earlier stages of sexual reproduction, particularly in understanding the molecular basis of pistil receptivity. The pistil is a complex collection of tissues that can be subdivided in three functionally distinct parts: the stigma, the style and the ovary (Fig. 1). Therefore, we next focused on determining the transcriptional profiles of these regions to identify differences in gene expression that might contribute to their function. Gradients of gene expression throughout the pistil have been previously documented. For example, the transcript of *STIGMA-SPECIFIC PROTEIN1* (*STIG1*), a small Cys-rich protein that promotes pollen tube growth, is expressed only in the stigma and upper style (Huang et al., 2014). A similar expression pattern was reported for the transcript of *MON9612* (Pitto et al., 2001). Based on this information and the visible morphological differences between the apical and basal portion of the style (particularly the presence trichomes in the basal portion), we decided to explore possible differences in gene expression between these subsections. Mature pistils were dissected as shown in Fig. 4A. To precisely identify female-specific or female-enriched genes, we included datasets representing male reproductive samples. We generated libraries from mature pollen grains in the bicellular stage (vegetative and generative cells present) and pollen tubes in the tricellular stage (vegetative and sperm cells present, after germination *in vitro* and incubation for 10 h). To accurately account for biological variability in our study when determining preferential gene expression, we generated four biological replicates per sample, with the exception of the stigma samples where three samples passed the RNA quality requirements (A260/280 between 1.8-2.1, A260/230 >1.5 and RIN value ≥ 7.6). We additionally included single-replicate transcriptomes from a range of Micro-Tom vegetative tissues and floral whorls published by Ezura et al., 2017, as well as transcriptomes from leaf, root and flower bud from the cultivar Heinz, deposited in the SolGenomics Network website (2 biological replicates each, Sato et al., 2012). Raw read quality was analyzed, overall the quality of our libraries was high (Phred score > 30). Sequencing adapter contamination was removed from the libraries if necessary. Transcript abundance was expressed in Transcripts per Million

(TPM) and, a gene was considered as expressed when TPM ≥ 0.1. After filtering based on this criterion, the pistil subsection samples retained ~70% of genes while pollen and pollen tube samples retained only 35-37% of genes. The latter observation is consistent with previous reports of a drastic reduction of the number of expressed genes in mature pollen in other species (*Arabidopsis*, rice; Honys and Twell, 2003; Wei et al, 2010).

The complexity of a transcriptome is determined by the number of genes contributing to the total bulk of transcription or total transcriptional output (TTO) (Carnici et al., 2000; Melé et al., 2015). In tissues with a low transcriptome complexity, few genes will account for the majority of the transcriptional output. We determined the fraction of the TTO for each gene in our datasets and found that pollen and pollen tubes have lower complexity transcriptomes relative to the pistil subsections with 51 and 44 genes accounting for ~50% of the TTO, respectively (Fig. 4B). In both cases, the functions of many of the highest expressed genes are related to cell wall metabolism (Sup. Table 4 and 5). Among the pistil samples, stigma had the lowest complexity with 473 genes accounting for half of the TTO. In this case, functions associated with the top expressed genes are lipid metabolism, defense and signal responses (Sup. Table 6). In the apical subsection of the style, 479 genes account for ~50% of the TTO; the basal portion of the style 526 genes and the same number of genes account for half of the bulk of transcription in the ovary (Fig. 4B). In the style and ovary subsections of the pistil the highest expressed gene belongs to the Defensin-like protein family (Sup. Table 7-9).

We then conducted a Principal Component Analysis (PCA) to visualize the variability amongst samples from different tissues (Fig. 4C). In our analysis, Principal Component 1 (PC1) and PC2 together explained 98% of the variation of gene expression among datasets, with 96% of variation explained by the first component (PC1). As expected from samples generated from the same tissue, our replicate samples displayed high similarity with respect to the two principal components. On the other hand, separation of the datasets along PC1 shows that the major differences in transcriptional makeup exist between male/female reproductive tissues. Pistil subsections are clustered together and in close distance with other floral whorls (sepal and petal) excluding anther, with the latter being closer to the cluster formed by the male gametophyte. In addition, distribution along PC2 shows a distinction between reproductive and vegetative tissues as previously reported (Ezura *et al*., 2017) (Fig. 4C).

Once we validated the quality of our transcriptomes, we identified genes with enriched expression in the pistil subsections and pollen and pollen tubes. Previous efforts have been done to identify ovary specific genes in tomato (Ezura et al., 2017; Pattison et al, 2015); however, to our knowledge, our study is the first effort to characterize the transcriptional profiles of subsections of the mature pistil. We performed a Differential Expression Analysis (DEA) by comparing each sample to the rest of the tissues, including libraries from other floral whorls (anther, sepal, petal) and vegetative tissues (leaf, root and flower bud). After DEA, we filtered the data and retained only genes with Log Fold Change (LFC) ≥ 2 and Q-value ≤ 0.05. We then looked for overlapping genes between the obtained lists to narrow down the number of genes with preferential expression per tissue/structure. After the second filtering, pollen grains had the greatest number of preferentially expressed genes (1229, Sup. Table 10), followed by pollen tubes (1126, Sup. Table 11). Within subsections of the pistil, the ovary displayed the highest number of differentially expressed genes (448, Sup. Table 15), followed by the stigma (411, Sup. Table 12), the apical portion of the style (57, Sup. Table 13) and lastly, the basal portion of the style (29, Sup. Table 14) (Fig. 4D).

In order to gain more insight on the functions of the genes identified in this study, we predicted the subcellular localization of the proteins they encode. Using a prediction tool based on machine learning (DeepLoc-01, Almagro Almenteros et al., 2017), we determined that in all tissues, the highest percent of genes were predicted to be nuclear or cytoplasmic (ranging from 16 to 26% for cytoplasmic and 17-30% for nuclear proteins, Sup. Fig. 3). Consistent with the highly secretory nature of the transmitting tract of Solanaceae species (Lush et al., 2000), the third most abundant localization in the apical (17%) and basal style (33%) is extracellular (Fig. 4E). Interestingly, in pollen and pollen tubes the third most abundantly predicted compartment is mitochondria. Although a direct relationship cannot be established solely on transcriptomic data, our findings are consistent with reports of a higher content of mitochondria in pollen relative to vegetative tissues in some species (Lee and Warmke, 1979) that presumably supports, at least in part, the highly energy-demanding process of pollen tube elongation (Selinski and Scheibe, 2014).

Characterization of expression profiles of the AGP family in tomato reproductive structures

In tomato, only a handful of AGPs had been characterized. To date, LeAGP1, a ubiquitously expressed classical AGP, is the biochemically and genetically best studied AGP (Showalter et al., 2000; Gao and Showalter et al., 2000; Sun et al., 2004). In addition, the study of Fragkostefanakis et al. (2012), identified 34 putative genes encoding AGPs, and further characterized the expression dynamics of *SlAGP2* (*Solyc04g074730.1*) and *SlAGP4* (*Solyc04g074730.1*) during fruit ripening and in response to wounding and hypoxia. Our immunological studies suggest an important role for AGPs in tomato reproduction, prior to fruit development (Figs. 2-3). Therefore, we were interested in investigating the expression of genes encoding AGPs in our datasets.

The AGP family is a highly diverse, multigenic group. They are highly glycosylated (up to 90% of their total weight corresponds to carbohydrates) and possess a high content of (hydroxy)proline (Hyp), alanine, serine and threonine in their sequences. AGP glycosylation occurs on hydroxyproline residues within clustered dipeptide motifs Ala-Hyp, Ser-Hyp, Val-Hyp, Gly-Hyp and Thr-Hyp (Ellis et al., 2010). Depending on the structure of their protein backbones, AGPs can be further divided into subgroups: classical AGPs, small AG-peptides, lysine-rich AGPs, hybrid AGPs (with glycosylation motifs belonging to different classes of Hydroxyproline-Rich Glycoproteins – HRGPs) and chimeric AGPs (with functionally and phylogenetically unrelated domains). Due to their biased amino acid composition and repetitive nature, several bioinformatics pipelines to predict the number of genes encoding AGPs have been developed (Schultz et al.,2002; Ma et al., 2017; Johnson et al., 2017, Showalter et al., 2010). In general, the available tools make use of different features in order to categorize a gene as member of the AGP family; for example, the percent of Pro, Ala, Ser and Thr present in the protein backbone, presence of a GPI anchor and signal peptide, sequence homology to known AGP genes, the presence of known chimeric domains in the sequence (e.g. Fasciclin-like domains, phytocyanin-like domains, xylogen-like domains, non-specific lipid transfer protein-like domains, early nodulin-like domains, etc.) or the presence and type of other glycosylation motifs for hybrid AGPs. As a consequence, the estimated number of sequences encoding for AGPs for a species varies depending on the prediction pipeline used. Thus, we compiled the ITAG identifiers of predicted sequences encoding for AGPs from three previously published reports (Ma et al., 2017; Johnson et al., 2017, Showalter et al., 2010). We then searched expression profiles among tomato tissues, with special interest in those displaying higher expression levels in our reproductive datasets (pistil subsections, pollen grains and tubes).

Out of the 161 presumed AGP genes in the tomato genome, 60 had detectable expression in our datasets and the vegetative datasets analyzed (Sup. Fig. 4). Among the genes represented in our analysis, the majority are classified as classical AGPs (16/60), followed by chimeric AGPs with Fasciclin-like domains (FLAs, 11/60). As anticipated, AGPs were expressed in a broad spectrum of tissues, both vegetative (especially root) and reproductive. Using hierarchical clustering of the AGP gene expression across tissues, we identified a cluster with enriched expression in reproductive tissues (Fig. 5A, Sup. Fig. 4, Sup. Fig. 5). This cluster included three genes with high expression in the pistil: a hybrid AGP-Extensin (*Solyc09g075580.1,* recently annotated as a short Extensin*, SlEXT14* by Ding et al., 2020*),* a Lysine-rich AGP (Lys-AGP, *Solyc07g052680.1,* annotated by Leszczuk et al., 2020 as *AGP19*) and a chimeric AGP with two Fasciclin-like domains (*Solyc07g065540.1,* annotated as *SlyFLA9* by Fragkostefanakis et al. 2012). The pistil expressed AGPs are represented in the Tomato Expression Atlas (TEA) in samples taken from dissected ovaries and in some cases through fruit development, although the expression levels reported are relatively low (Sup. Fig. 6A). In this cluster we also observed a classical AGP, *Solyc05g049890.1,* expressed in male reproductive structures (pollen grains and tubes). This AGP was represented in the Plant eFP data, as expressed in unopened flowers, possibly due to the lack of more refined datasets for floral parts other than ovaries (Sup. Fig. 6B). Since the data we analyzed was derived from both the Micro-Tom and Heinz 1706 cultivars, we compared the coding sequences of the reproductive AGPs in these two backgrounds and found that the predicted coding sequences were identical except for *Solyc09g075580.1* which possessed two amino acid substitutions (F15V, L66S) and an indel mutation resulting in an additional proline residue within an extensin $S(P)_{3-5}$ motif in the Micro-Tom sequence (Sup. Fig. 7). As a complement to the identification of AGPs, we also searched for tomato genes encoding hydroxyproline galactosyltransferases (hyp-GALT and HPGT) proteins. These enzymes are required for the initiation of AGP glycosylation and their presence in the tomato genome is predicted by the detection of AGP glycans (Basu et al., 2013; Ogawa-Ohnishi and Matsubayashi, 2015). We identified six tomato homologues of the Arabidopsis *GALT* and *HPGT* sequences and found they were broadly expressed in the pistil and vegetative tissue with relatively weaker expression in the pollen samples (Sup. Fig. 8).

From the pistil expressed AGPs, we selected those genes with the highest expression values (mean TPM) for further validation with an independent approach. The expression patterns of *Solyc09g075580.1* (hereafter *SlyHAE*) and *SlyFLA9* were confirmed by quantitative RT-PCR. *SlyHAE* displayed a higher level of expression in the basal portion of the style (Fig. 5C). The

protein sequence encoded by *SlyHAE* has relatively high content of Pro (28%), six predicted AGP glycomodules and two glycosylation motifs found in a different class of HRGPs, Extensins (Fig. 5B). AGP-EXT hybrids like the class III Pistil Extensin-like Protein (PELPIII) have been reported in other members of the Solanaceae family like tobacco (Eberle et al., 2013; de Graaf et al., 2003); however, *SlyHAE* shares poor sequence similarity to this gene family (< 12% sequence similarity to *N. tobaccum* PELPIII). Likewise, alignment of *SlyHAE* to other AGPs in tobacco known to be involved in reproduction (TTS, 120 KDa) displayed poor sequence similarity (supplemental table 16). Two genes encoded in the genomes of *Nicotiana benthamiana* (*Niben101Scf01623g08001.1*) and *Solanum melongena* (*SMEL_009g330490.1.01*) displayed high sequence similarity to *SlyHAE* (61.48% and 85.61%, respectively) (Sup. Fig. 9). No known functions for these genes in tobacco or eggplant have been reported. Outside of the Solanaceae family, no homologs of *SlyHAE* were identified.

SlyFLA9, displayed a high level of expression in the ovary (Fig. 5C). *SlyFLA9* is a chimeric AGP with two predicted Fasciclin domains (FAS), a C-terminus GPI anchor and 20 putative AGP glycomodules (Fig. 5B). The FAS domains are conserved in Eukarya and are predicted to participate in cell-cell adhesion by an unknown mechanism (Seifert, 2018). Sequence alignment of *SlyFLA9* produces significant hits to *FLA1* in *Arabidopsis thaliana*, which also possesses 2 FAS domains (72.9% similarity). The loss of function of *FLA1* in Arabidopsis displayed shoot regeneration defects and, based on its expression pattern, is presumed to participate during embryogenesis and seed development (Johnson et al., 2011).

Based on the high expression level of *SlyFLA9* observed in the ovary (Fig. 5B), we hypothesized it may contribute to the AGP glycan signals we observed in this organ (Fig. 1B, Fig. 2 and Fig. 3). To evaluate the spatial distribution of *SlyFLA9* transcripts, we generated DIG-labeled RNA probes for *in situ* detection in ovaries dissected from mature, unpollinated pistils. We observed high accumulation of the *SlyFLA9* transcript in the pericarp of the ovary and the outer integuments of individual ovules (Fig. 5D). Based on the protein structure of *SlyFLA9* and its expression pattern, our results suggest a role during ovary maturation and ovule receptivity.

Discussion

The interaction between pollen and pistil is an ephemeral event during a plant's life cycle, yet the proper establishment of their molecular dialogue is crucial for successful fertilization. The

AGP family has been implicated in virtually all steps of pollen-pistil interaction, paving the pathway along the sporophytic tissues of the pistil towards a receptive ovule (Dresselhaus and Franklin-Tong, 2013; Su and Higashiyama, 2018). Here, we report the distribution of AG glycans throughout pistil development and following fertilization (Figs. 1, 2 and 3). Our results show that AGPs progressively accumulate in the stigmatic surface of the pistil, consistent with reports in other species with wet stigmas (Losada and Herrero, 2012; Costa et al., 2013; Gell et al., 1986) and suggesting that AGPs present in stigmatic exudates play a role for the earlier stages of pollen-pistil interaction.

After pollen tubes penetrate the stigmatic cells, they elongate through the intercellular spaces of the transmitting tract in species with solid styles (Gotelli et al., 2017). Once in the style, directional pollen tube growth is thought to be regulated by chemical cues (Cheung et al. 1995) in combination with mechanical forces (Reinmann et al., 2020). AGPs in the transmitting tract of solanaceous species have been implicated in stylar guidance: in tobacco*,* TTS, a member of the AGP family, displays a basipetal glycosylation gradient along the style of tobacco pistils, where it acts as a pollen tube attractant and as a source of nutrients upon glycan hydrolysis (Wu et al., 1995). Interestingly, genetic studies in the same species demonstrated that, in contrast to species with dry stigmas where incompatible pollination blockage occurs in the papillae (Fujii et al., 2019), the transmitting tract in the style acts as pre-zygotic barrier for interspecific pollination (Goldman et al., 1994). Several studies have highlighted the importance of the hybrid Extensin-AGP, PELPIII, during interspecies incompatibility in tobacco. PELPIII accumulates in the transmitting tract of the style and, after pollination, is translocated to the cell wall of the growing pollen tubes (de Graaf et al., 2003), inhibiting interspecific pollen tube growth (Alves et al., 2019). Our histological studies showed that β-Yariv reactive AGPs accumulate in the transmitting tract of the style of mature pistils (Fig. 1B), potentially implicating AGPs in stylar guidance.

Due to the thickness and size of the pistil of solanaceous species, little information is available regarding the passage of pollen tubes from the transmitting tract to the ovary and how exactly pollen tubes respond to ovular cues once inside. In a study performed by Lush et al. (2000) in *Nicotiana alata* ovaries, which enclose ~400 ovules, pollen tubes expressing a GUS reporter appear to fertilize ovules in a random fashion. In addition, they reported pollen tube ‘meandering’ within the ovary, where pollen tubes exited the placental surface, bypassing ovules and growing through their external surface in a disorganized manner. We also observed

this behavior in tomato ovaries 24 h after pollination when we removed the pericarp to expose the ovules and growing pollen tubes (Sup. Fig. 1). Despite this seemingly random pollen tube behavior, they are capable of finding receptive ovules and fertilizing them, suggesting that a guidance mechanism indeed exists in this species. The micropyle, formed by the integuments that surround the embryo sac, is the pollen tube's entry point to the ovule. The study of mutants with anomalous integuments revealed partial defects in pollen tube ovule targeting, suggesting a role for the micropyle in guidance (Lora et al., 2018). Although how the micropyle participates in pollen guidance is still unknown, the accumulation of AGPs in this region (Fig. 3; Coimbra et al., 2018; Hou et al., 2016; Losada and Herrero, 2019) suggest potential functions in this process. Additional evidence of AGP involvement in ovular pollen guidance comes from the discovery of AMOR, an arabinogalactan derived sugar produced by the sporophytic tissue of the ovule that renders pollen tubes responsive to synergid signaling peptides (Mizukami et al., 2016). Whether a pollen receptor that binds to AMOR exists, remains to be determined. It has been suggested that given the importance of Ca^{2+} dynamics for pollen tube growth and the ability of AGPs to bind Ca^{2+} in a pH-dependent manner, AGPs might influence cell growth by acting as a periplasmic Ca^{2+}reservoir (Lamport and Várnai, 2013) or in the case of AMOR, a diffusible Ca^{2+} carrier (Lamport et al., 2018).

AGPs had been implicated in additional developmental processes, here we also report a reorganization of the AGP epitopes in the ovules after fertilization (Fig. 2, 3), implicating AGPs in latter steps of embryo and/or fruit development. The latter has been reported for fruit ripening and softening (Leszczuk et al., 2019).

Immunological studies have proven useful to determining spatial regulation of AGPs in a number of species; however, for the most part, the identity of their protein backbones remains to be characterized. Several efforts combining amino acid bias quantification, prediction of signal peptides and homology searches have predicted members of the AGP family in different species (Ma et al., 2017; Johnson et al., 2017; Showalter et al., 2010). We compiled the tomato AGP sequences from these datasets; these included sequences corresponding to classical AGPs, hybrid and chimeric AGPs and small peptides (Fig. 5 and Sup. Fig. 4). To evaluate their expression profiles across reproductive tissues, we first generated a set of RNA-seq libraries derived from subsections of the mature pistil and pollen grains and tubes (Fig. 4). From the original gene list of 161, we detected expression of 60 genes. Consistent with comprehensive studies in other species like *Arabidopsis* (Pereira et al., 2014), the AGP family displays different

levels of expression across plant organs. Classical AGPs and FLAs were the most abundantly expressed subclasses of AGPs in tomato tissues (sup. Fig. 4). We identified and validated expression of two AGPs: a hybrid AGP-Extensin, *SlyHAE* (*Solyc09g075580.1*) with no homologs outside the Solanaceae family and with high expression in the style (Fig. 5) and a Fasciclin-like AGP, *SlyFLA9* (*Solyc07g065540.1*) with high expression levels in the ovary, particularly in the outer cell layer of the integuments and the pericarp (Fig. 5C and D). Fasciclin-like AGPs have been implicated in the development of pollen in *Arabidopsis* (Li et al., 2010) and fiber development in cotton (Huang et al., 2013). The exact mechanism of action of FLAs is poorly understood. It has been hypothesized that due to the presence of the Fas1 domain in their protein backbones, plant FLAs might function like those reported in metazoans. Drosophila Fas1 proteins function as cell-cell adhesion molecules during neuronal development (Zhong and Shanley, 1995). The structure of the Fas1 domain has been crystallized; however, the mechanism of action remains to be elucidated (Twarda-Clapa et al., 2018). The majority of FLAs have predicted GPI-anchors and it has been hypothesized that they can act as diffusible signaling molecules upon phospholipase mediated cleavage (Schultz et al., 1998). Cell-cell adhesion mediated translocation has been proposed as a mode of pollen tube guidance inside the ovary (Lush et al., 2000). The expression pattern of *SlyFLA9* makes it an interesting candidate for future genetic studies to evaluate its role during pollen-pistil interaction.

Although efforts to identify pistil-expressed genes have been published elsewhere (Ezura et al., 2017), our experimental design allowed us to identify the transcriptional profiles corresponding to stigma, apical and basal style and ovary. After differential expression analyses, careful inspection of the identified genes highlighted the potential for widespread usage of our datasets as tools for further characterization of genes upregulated in each pistil subsection. For example, in the stigma libraries, we found significant enrichment of genes related to lipid metabolism (Sup. Tables 2 and 3). Chemical composition studies of stigma exudate in other members of the Solanaceae family, like tobacco and petunia, revealed that it is lipid-rich. These lipids play an important role in pollen hydration, germination and directing pollen growth towards the aqueous phase underneath the hydrophobic exudate (Lush et al., 2000; Konar and Linskens, 1966; Wolters-Arts et al., 1998). Several WSD1-like O-acyltransferases genes (*Solyc03g083385.1*, *Solyc03g083380.3*, *Solyc01g095930.3* and *Solyc10g009430.3*) had enriched expression in the stigma (LFC > 3). *WSD1* genes in *Arabidopsis* are involved in cuticle biosynthesis (Li et al., 2008). The presence of a cuticle layer in species with wet stigma has been documented either

as a continuous layer (in *Vitis vinifera*, Ciampolini et al., 1996) or as a discontinuous layer, ruptured by the secretion of exudates (in *Petunia hybrida*, Konar and Linskens, 1966).

In both portions of the style, genes with predicted extracellular localization were the third most enriched category (Fig. 4D). Consistent with the literature, our apical style library detected enriched expression of genes annotated as STIG1 proteins (*Solyc03g120960.1* and *Solyc03g120955.1*) and the *MON9612* transcript (*Solyc02g093580.*3; Huang et al., 2014; Pitto et al., 2001). In this subsection, we also identified several enriched genes with predicted activities related to carbohydrate metabolism and cell wall polymer remodeling such as cell wall invertases (*Solyc01g088590.4, Solyc03g121680.2*), glycosyltransferases (*Solyc08g077080.1, Solyc05g012670.1*), pectinesterase (*Solyc01g099960*.3), polygalacturonase (*Solyc12g019180.2*) and α/ β-hydrolases (*Solyc08g083190.3*). Both, carbohydrate metabolism and pectin remodeling have recently been implicated in pollen-pistil interaction. In tomato pistils, cell wall invertases (*CWIN*) and hexose transporters are upregulated in response to pollination, in the style briefly after pollination to presumably support pollen tube growth, and in the ovary two days after pollination, potentially induced after fertilization (Shen et al., 2019). Pistil expressed polygalacturonases (PGs) are pectin hydrolytic enzymes that have been recently implicated in interspecific incompatibility in tobacco, where *in vitro* and semi-*in vivo* assays showed that PG activity inhibits pollen tube growth (Liao et al., 2020).

In our study we identified several genes encoding peroxidases, with expression in the stigma (*Solyc10g075120.2*), four genes in the basal style (*Solyc10g047110.2, Solyc05g052280.3, Solyc03g025380.3, Solyc05g046020.3*) and four genes in the ovary (*Solyc10g078890.2, Solyc09g072700.3, Solyc01g104860.3, Solyc02g084790.3*). Peroxidase activity is commonly used as a test for stigmatic receptivity (Dafni and Maués, 1998). Peroxidase activity and accumulation of Reactive Oxygen Species (ROS, specifically H_2O_2) has been described in the stigmas of several species (McInnis et al., 2017). In tobacco, peroxidase activity measured along the pistil is distributed as a gradient, with the highest activity in the stigma, decreasing activity in the style and a slight increase of activity in the ovary with distinct peroxidase isozymes in each subsection (Bredemeijer, 1984). ROS signaling has been linked to pollen germination and hydration in the stigma, to pollen tube elongation in the transmitting tract and to ovular guidance in the ovary (Zhang et al., 2020).

Small secreted Cysteine-rich proteins (CRPs) play an important role during ovular guidance. CRPs are subclassified into three groups: Defense-like proteins (DEFL), Early Culture Abundant 1 proteins (ECA1) and nonspecific Lipid Transfer proteins (Sprunk et al., 2014). DEFL genes encode small peptides, some of which have species-specific pollen attractant functions (LURE peptides; Takeuchi and Higashiyama, 2012; Higashiyama and Yang, 2017). In *Arabidopsis* a member of the ECA1 subgroup, Egg Cell 1 (EC1) accumulates in the egg cell and promotes sperm cell fusion during fertilization (Sprunk et al., 2012). In our analysis, we identified several CRPs highly expressed in the ovary: four predicted ECA1 genes (*Solyc05g010190.1, Solyc06g048400.1, Solyc11g008725.1* and *Solyc07g006210.1,* LFC > 4.7, the first gene was also identified by the study of Ezura et al., 2017) and three predicted DEFL genes (*Solyc09g074440.3, Solyc06g075200.1* and *Solyc07g017570.2*; LFC >3.3).

Our study identified, in addition to pistil expressed AGPs, a number of genes with potential functions in the establishment of a receptive pistil in tomato. The predicted functions of the identified genes cover a broad spectrum: from cell wall modifying enzymes to small secreted signaling peptides. Functional characterization of these genes will allow us to further understand the basis of pistil receptivity and by performing comparative studies, understand how these genes involved in sexual reproduction evolved.

Declarations

Funding:
This work was supported by the National Science Foundation under Grant No. IOS-1755482 to CAM. CMLM receives fellowship funding from the Mexican Council of Science and Technology (CONACYT) (773973).

Conflicts of interest:
The authors declare no conflict of interest.

Availability of data and material:
Data will be deposited in the NCBI database.

Authors' contributions:

CMLM designed and performed experiments, analyzed data and wrote the manuscript.

CAM conceived the study, designed experiments and edited the manuscript.

References

Acosta-García, G., & Vielle-Calzada, J.-P. (2004). A classical arabinogalactan protein is essential for the initiation of female gametogenesis in Arabidopsis. *The Plant Cell*, *16*(10), 2614–2628. https://doi.org/10.1105/tpc.104.024588

Almagro Armenteros, J. J., Sønderby, C. K., Sønderby, S. K., Nielsen, H., & Winther, O. (2017). DeepLoc: Prediction of protein subcellular localization using deep learning. *Bioinformatics*, *33*(21), 3387–3395. https://doi.org/10.1093/bioinformatics/btx431

Alves, C. M. L., Noyszewski, A. K., & Smith, A. G. (2019). Structure and function of class III pistil-specific extensin-like protein in interspecific reproductive barriers. *BMC Plant Biology*, *19*(1), 118. https://doi.org/10.1186/s12870-019-1728-8

Basu, D., Liang, Y., Liu, X., Himmeldirk, K., Faik, A., Kieliszewski, M., Held, M. & Showalter AM. (2013). Functional identification of a hydroxyproline-o-galactosyltransferase specific for arabinogalactan protein biosynthesis in Arabidopsis. *J Biol Chem.* 288(14):10132-43. doi: 10.1074/jbc.M112.432609

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525–527. https://doi.org/10.1038/nbt.3519

Bredemeijer, G. M. M. (1984). The role of peroxidases in pistil-pollen interactions. *Theoretical and Applied Genetics*, *68*(3), 193–206. https://doi.org/10.1007/BF00266889

Busby, M. A., Stewart, C., Miller, C. A., Grzeda, K. R., & Marth, G. T. (2013). Scotty: A web tool for designing RNA-Seq experiments to measure differential gene expression. *Bioinformatics*, *29*(5), 656–657. https://doi.org/10.1093/bioinformatics/btt015

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., & Hayashizaki, Y. (2000). Normalization and Subtraction of Cap-Trapper-Selected cDNAs to Prepare Full-Length cDNA Libraries for Rapid Discovery of New Genes. *Genome Research*, *10*(10), 1617–1630. https://doi.org/10.1101/gr.145100

Challinor, A. J., Watson, J., Lobell, D. B., Howden, S. M., Smith, D. R., & Chhetri, N. (2014). A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change*, *4*(4), 287–291. https://doi.org/10.1038/nclimate2153

Cheung, A. Y., Boavida, L. C., Aggarwal, M., Wu, H.-M., & Feijó, J. A. (2010). The pollen tube journey in the pistil and imaging the in vivo process by two-photon microscopy. *Journal of Experimental Botany*, *61*(7), 1907–1915. https://doi.org/10.1093/jxb/erq062

Cheung, A. Y., Wang, H., & Wu, H. (1995). A floral transmitting tissue-specific glycoprotein attracts pollen tubes and stimulates their growth. *Cell*, *82*(3), 383–393. https://doi.org/10.1016/0092-8674(95)90427-1

Ciampolini, F., Faleri, C., Di Pietro, D., & Cresti, M. (1996). Structural and Cytochemical Characteristics of the Stigma and Style in Vitis vinifera L. var. Sangiovese (Vitaceae). *Annals of Botany*, *78*(6), 759–764.

Coimbra, S., Almeida, J., Junqueira, V., Costa, M. L., & Pereira, L. G. (2007). Arabinogalactan proteins as molecular markers in Arabidopsis thaliana sexual reproduction. *Journal of Experimental Botany*, *58*(15–16), 4027–4035. https://doi.org/10.1093/jxb/erm259

Costa, M., Pereira, A. M., Rudall, P. J., & Coimbra, S. (2013). Immunolocalization of arabinogalactan proteins (AGPs) in reproductive structures of an early-divergent angiosperm, Trithuria (Hydatellaceae). *Annals of Botany*, *111*(2), 183–190. https://doi.org/10.1093/aob/mcs256

Covey, P. A., Subbaiah, C. C., Parsons, R. L., Pearce, G., Lay, F. T., Anderson, M. A., Ryan, C. A., & Bedinger, P. A. (2010). A Pollen-Specific RALF from Tomato That Regulates Pollen Tube Elongation. *Plant Physiology*, *153*(2), 703–715. https://doi.org/10.1104/pp.110.155457

da Costa, M. L., Lopes, A. L., Amorim, M. I., & Coimbra, S. (2017). Immunolocalization of AGPs and Pectins in Quercus suber Gametophytic Structures. In A. Schmidt (Ed.), *Plant Germline Development: Methods and Protocols* (pp. 117–137). Springer. https://doi.org/10.1007/978-1-4939-7286-9_11

Dafni, A., & Maués, M. M. (1998). A rapid and simple procedure to determine stigma receptivity. *Sexual Plant Reproduction*, *11*(3), 177–180. https://doi.org/10.1007/s004970050138

de Graaf, B. H. J., Knuiman, B. A., Derksen, J., & Mariani, C. (2003). Characterization and localization of the transmitting tissue-specific PELPIII proteins of Nicotiana tabacum. *Journal of Experimental Botany*, *54*(380), 55–63. https://doi.org/10.1093/jxb/erg002

Ding, Q., Yang, X., Pi, Y., Li, Z., Xue, J., Chen, H., Li, Y., & Wu, H. (2020). Genome-wide identification and expression analysis of extensin genes in tomato. Genomics, 112(6), 4348–4360. https://doi.org/10.1016/j.ygeno.2020.07.029

Dresselhaus, T., & Franklin-Tong, N. (2013). Male–Female Crosstalk during Pollen Germination, Tube Growth and Guidance, and Double Fertilization. *Molecular Plant*, *6*(4), 1018–1036. https://doi.org/10.1093/mp/sst061

Eberle, C. A., Anderson, N. O., Clasen, B. M., Hegeman, A. D., & Smith, A. G. (2013). PELPIII: The class III pistil-specific extensin-like Nicotiana tabacum proteins are essential for interspecific incompatibility. *The Plant Journal*, *74*(5), 805–814. https://doi.org/10.1111/tpj.12163

Edlund, A. F., Swanson, R., & Preuss, D. (2004). Pollen and Stigma Structure and Function: The Role of Diversity in Pollination. *The Plant Cell*, *16*(suppl 1), S84–S97. https://doi.org/10.1105/tpc.015800

Ellis, M., Egelund, J., Schultz, C. J., & Bacic, A. (2010). Arabinogalactan-Proteins: Key Regulators at the Cell Surface? *Plant Physiology*, *153*(2), 403–419. https://doi.org/10.1104/pp.110.156000

Ezura, K., Ji-Seong, K., Mori, K., Suzuki, Y., Kuhara, S., Ariizumi, T., & Ezura, H. (2017). Genome-wide identification of pistil-specific genes expressed during fruit set initiation in tomato (Solanum lycopersicum). *PLOS ONE*, *12*(7), e0180003. https://doi.org/10.1371/journal.pone.0180003

Fragkostefanakis, S., Dandachi, F., & Kalaitzis, P. (2012). Expression of arabinogalactan proteins during tomato fruit ripening and in response to mechanical wounding, hypoxia and anoxia. *Plant Physiology and Biochemistry: PPB*, *52*, 112–118. https://doi.org/10.1016/j.plaphy.2011.12.001

Fujii, S., Tsuchimatsu, T., Kimura, Y., Ishida, S., Tangpranomkorn, S., Shimosato-Asano, H., Iwano, M., Furukawa, S., Itoyama, W., Wada, Y., Shimizu, K. K., & Takayama, S. (2019). A stigmatic gene confers interspecies incompatibility in the Brassicaceae. *Nature Plants*, *5*(7), 731–741. https://doi.org/10.1038/s41477-019-0444-6

Gao, M., & Showalter, A. M. (2000). Immunolocalization of LeAGP-1, a modular arabinogalactan-protein, reveals its developmentally regulated expression in tomato. *Planta*, *210*(6), 865–874. https://doi.org/10.1007/s004250050691

Gell, A. C., Bacic, A., & Clarke, A. E. (1986). Arabinogalactan-Proteins of the Female Sexual Tissue of Nicotiana alata: I. Changes during Flower Development and Pollination. *Plant Physiology*, *82*(4), 885–889. https://doi.org/10.1104/pp.82.4.885

Goldman, M. h., Goldberg, R. b., & Mariani, C. (1994). Female sterile tobacco plants are produced by stigma-specific cell ablation. *The EMBO Journal*, *13*(13), 2976–2984. https://doi.org/10.1002/j.1460-2075.1994.tb06596.x

Gotelli, M. M., Lattar, E. C., Zini, L. M., & Galati, B. G. (2017). Style morphology and pollen tube pathway. *Plant Reproduction*, *30*(4), 155–170. https://doi.org/10.1007/s00497-017-0312-3

Heslop-Harrison, Y., & Shivanna, K. R. (1977). The Receptive Surface of the Angiosperm Stigma. *Annals of Botany*, *41*(6), 1233–1258. https://doi.org/10.1093/oxfordjournals.aob.a085414

Higashiyama, T., & Takeuchi, H. (2015). The Mechanism and Key Molecules Involved in Pollen Tube Guidance. *Annual Review of Plant Biology*, *66*(1), 393–413. https://doi.org/10.1146/annurev-arplant-043014-115635

Higashiyama, T., & Yang, W. (2017). Gametophytic Pollen Tube Guidance: Attractant Peptides, Gametic Controls, and Receptors1. *Plant Physiology*, *173*(1), 112–121. https://doi.org/10.1104/pp.16.01571

Honys, D., & Twell, D. (2004). Transcriptome analysis of haploid male gametophyte development in Arabidopsis. *Genome Biology*, *5*(11), R85. https://doi.org/10.1186/gb-2004-5-11-r85

Hou, Y., Guo, X., Cyprys, P., Zhang, Y., Bleckmann, A., Cai, L., Huang, Q., Luo, Y., Gu, H., Dresselhaus, T., Dong, J., & Qu, L.-J. (2016). Maternal ENODLs Are Required for Pollen Tube Reception in Arabidopsis. *Current Biology*, *26*(17), 2343–2350. https://doi.org/10.1016/j.cub.2016.06.053

Huang, G.-Q., Gong, S.-Y., Xu, W.-L., Li, W., Li, P., Zhang, C.-J., Li, D.-D., Zheng, Y., Li, F.-G., & Li, X.-B. (2013). A Fasciclin-Like Arabinogalactan Protein, GhFLA1, Is Involved in Fiber Initiation and Elongation of Cotton. *Plant Physiology*, *161*(3), 1278–1290. https://doi.org/10.1104/pp.112.203760

Huang, W.-J., Liu, H.-K., McCormick, S., & Tang, W.-H. (2014). Tomato Pistil Factor STIG1 Promotes in Vivo Pollen Tube Growth by Binding to Phosphatidylinositol 3-Phosphate and the

Extracellular Domain of the Pollen Receptor Kinase LePRK2. *The Plant Cell*, *26*(6), 2505–2523. https://doi.org/10.1105/tpc.114.123281

Javelle, M., Marco, C. F., & Timmermans, M. (2011). In Situ Hybridization for the Precise Localization of Transcripts in Plants. *JoVE (Journal of Visualized Experiments)*, *57*, e3328. https://doi.org/10.3791/3328

Johnson, K. L., Cassin, A. M., Lonsdale, A., Bacic, A., Doblin, M. S., & Schultz, C. J. (2017). Pipeline to Identify Hydroxyproline-Rich Glycoproteins. *Plant Physiology*, *174*(2), 886–903. https://doi.org/10.1104/pp.17.00294

Johnson, K. L., Kibble, N. A. J., Bacic, A., & Schultz, C. J. (2011). A Fasciclin-Like Arabinogalactan-Protein (FLA) Mutant of Arabidopsis thaliana, fla1, Shows Defects in Shoot Regeneration. *PLoS ONE*, *6*(9). https://doi.org/10.1371/journal.pone.0025154

Johnson, M. A., & Preuss, D. (2002). Plotting a Course: Multiple Signals Guide Pollen Tubes to Their Targets. *Developmental Cell*, *2*(3), 273–281. https://doi.org/10.1016/S1534-5807(02)00130-2

Kandasamy, M. K., Nasrallah, J. B., & Nasrallah, M. E. (1994). Pollen-pistil interactions and developmental regulation of pollen tube growth in Arabidopsis. *Development*, *120*(12), 3405–3418.

Kimura, S., & Sinha, N. (2008). Tomato (Solanum lycopersicum): A Model Fruit-Bearing Crop. *Cold Spring Harbor Protocols*, *2008*(11), pdb.emo105. https://doi.org/10.1101/pdb.emo105

Knox, J. P., Linstead, P. J., Cooper, J. P., C., & Roberts, K. (1991). Developmentally regulated epitopes of cell surface arabinogalactan proteins and their relation to root tissue pattern formation. *The Plant Journal*, *1*(3), 317–326. https://doi.org/10.1046/j.1365-313X.1991.t01-9-00999.x

Knox, J. Paul. (1995). Developmentally regulated proteoglycans and glycoproteins of the plant cell surface. *The FASEB Journal*, *9*(11), 1004–1012. https://doi.org/10.1096/fasebj.9.11.7544308

Konar, R. N., & Linskens, H. F. (1966). THE MORPHOLOGY AND ANATOMY OF THE STIGMA OF PETUNIA HYBRIDA. *Planta*, *71*(4), 356–371.

Labarca, C., & Loewus, F. (1972). The Nutritional Role of Pistil Exudate in Pollen Tube Wall Formation in Lilium longiflorum: I. Utilization of Injected Stigmatic Exudate. *Plant Physiology*, *50*(1), 7–14. https://doi.org/10.1104/pp.50.1.7

Lamport, D. T. A., & Várnai, P. (2013). Periplasmic arabinogalactan glycoproteins act as a calcium capacitor that regulates plant growth and development. *New Phytologist*, *197*(1), 58–64. https://doi.org/10.1111/nph.12005

Lamport, D. T. A., Tan, L., Held, M. A., & Kieliszewski, M. J. (2018). Pollen tube growth and guidance: Occam's razor sharpened on a molecular arabinogalactan glycoprotein Rosetta Stone. *The New Phytologist*, *217*(2), 491–500. https://doi.org/10.1111/nph.14845

Lee, S.-L. J., & Warmke, H. E. (1979). Organelle Size and Number in Fertile and T-Cytoplasmic Male-Sterile Corn. *American Journal of Botany*, *66*(2), 141–148. https://doi.org/10.1002/j.1537-2197.1979.tb06206.x

Lennon, K. A., Roy, S., Hepler, P. K., & Lord, E. M. (1998). The structure of the transmitting tissue of Arabidopsis thaliana (L.) and the path of pollen tube growth. *Sexual Plant Reproduction*, *11*(1), 49–59. https://doi.org/10.1007/s004970050120

Leszczuk, A., Chylińska, M., & Zdunek, A. (2019). Enzymes and vitamin C as factors influencing the presence of arabinogalactan proteins (AGPs) in Solanum lycopersicum fruit. *Plant Physiology and Biochemistry*, *139*, 681–690. https://doi.org/10.1016/j.plaphy.2019.04.035

Leszczuk, A., Kalaitzis, P., Blazakis, K. N., & Zdunek, A. (2020). The role of arabinogalactan proteins (AGPs) in fruit ripening—A review. *Horticulture Research*, *7*. https://doi.org/10.1038/s41438-020-00397-8

Li, F., Wu, X., Lam, P., Bird, D., Zheng, H., Samuels, L., Jetter, R., & Kunst, L. (2008). Identification of the Wax Ester Synthase/Acyl-Coenzyme A:Diacylglycerol Acyltransferase WSD1 Required for Stem Wax Ester Biosynthesis in Arabidopsis. *Plant Physiology*, *148*(1), 97–107. https://doi.org/10.1104/pp.108.123471

Li, J., Yu, M., Geng, L.-L., & Zhao, J. (2010). The fasciclin-like arabinogalactan protein gene, FLA3, is involved in microspore development of Arabidopsis. *The Plant Journal*, *64*(3), 482–497. https://doi.org/10.1111/j.1365-313X.2010.04344.x

Liao, J., Chen, Z., Wei, X., Tao, K., Zhang, J., Qin, X., Pan, Z., Ma, W., Pan, L., Yang, S., Wang, M., Ou, X., & Chen, S. (2020). Identification of pollen and pistil polygalacturonases in Nicotiana tabacum and their function in interspecific stigma compatibility. *Plant Reproduction*, *33*(3), 173–190. https://doi.org/10.1007/s00497-020-00393-x

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods*, *25*(4), 402–408. https://doi.org/10.1006/meth.2001.1262

Lopes, A. L., Costa, M. L., Sobral, R., Costa, M. M., Amorim, M. I., & Coimbra, S. (2016). Arabinogalactan proteins and pectin distribution during female gametogenesis in Quercus suber L. *Annals of Botany*, *117*(6), 949–961. https://doi.org/10.1093/aob/mcw019

Lora, J., Laux, T., & Hormaza, J. I. (2019). The role of the integuments in pollen tube guidance in flowering plants. *New Phytologist*, *221*(2), 1074–1089. https://doi.org/10.1111/nph.15420

Lord, E. M., & Heslop-Harrison, Y. (1984). Pollen-Stigma Interaction in the Leguminosae: Stigma Organization and the Breeding System in Vicia faba L. *Annals of Botany*, *54*(6), 827–836.

Losada, J. M., & Herrero, M. (2012). Arabinogalactan-protein secretion is associated with the acquisition of stigmatic receptivity in the apple flower. *Annals of Botany*, *110*(3), 573–584. https://doi.org/10.1093/aob/mcs116

Losada, J. M., & Herrero, M. (2017). Pollen tube access to the ovule is mediated by glycoprotein secretion on the obturator of apple (Malus × domestica, Borkh). *Annals of Botany*, *119*(6), 989–1000. https://doi.org/10.1093/aob/mcw276

Losada, J. M., Herrero, M., Hormaza, J. I., & Friedman, W. E. (2014). Arabinogalactan proteins mark stigmatic receptivity in the protogynous flowers of Magnolia virginiana (Magnoliaceae). *American Journal of Botany*, *101*(11), 1963–1975. https://doi.org/10.3732/ajb.1400280

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lush, W. M., Spurck, T., & Joosten, R. (2000). Pollen Tube Guidance by the Pistil of a Solanaceous Plant. *Annals of Botany*, *85*(suppl_1), 39–47. https://doi.org/10.1006/anbo.1999.1059

Ma, Y., Yan, C., Li, H., Wu, W., Liu, Y., Wang, Y., Chen, Q., & Ma, H. (2017). Bioinformatics Prediction and Evolution Analysis of Arabinogalactan Proteins in the Plant Kingdom. *Frontiers in Plant Science*, *8*. https://doi.org/10.3389/fpls.2017.00066

MacAlister, C. A., Park, S. J., Jiang, K., Marcel, F., Bendahmane, A., Izkovich, Y., Eshed, Y., & Lippman, Z. B. (2012). Synchronization of the flowering transition by the tomato TERMINATING FLOWER gene. *Nature Genetics*, *44*(12), 1393–1398. https://doi.org/10.1038/ng.2465

Mareri, L., Faleri, C., Romi, M., Mariani, C., Cresti, M., & Cai, G. (2016). Heat stress affects the distribution of JIM8-labelled arabinogalactan proteins in pistils of Solanum lycopersicum cv Micro-Tom. *Acta Physiologiae Plantarum*, *38*(7),184. https://doi.org/10.1007/s11738-016-2203-x

McInnis, S. M., Desikan, R., Hancock, J. T., & Hiscock, S. J. (2006). Production of reactive oxygen species and reactive nitrogen species by angiosperm stigmas and pollen: Potential signalling crosstalk? *New Phytologist*, *172*(2), 221–228. https://doi.org/10.1111/j.1469-8137.2006.01875.x

Melé, M., Ferreira, P. G., Reverter, F., DeLuca, D. S., Monlong, J., Sammeth, M., Young, T. R., Goldmann, J. M., Pervouchine, D. D., Sullivan, T. J., Johnson, R., Segrè, A. V., Djebali, S.,

Niarchou, A., Consortium, T. Gte., Wright, F. A., Lappalainen, T., Calvo, M., Getz, G., … Guigó, R. (2015). The human transcriptome across tissues and individuals. *Science*, *348*(6235), 660–665. https://doi.org/10.1126/science.aaa0355

Mizukami, A. G., Inatsugi, R., Jiao, J., Kotake, T., Kuwata, K., Ootani, K., Okuda, S., Sankaranarayanan, S., Sato, Y., Maruyama, D., Iwai, H., Garénaux, E., Sato, C., Kitajima, K., Tsumuraya, Y., Mori, H., Yamaguchi, J., Itami, K., Sasaki, N., & Higashiyama, T. (2016). The AMOR Arabinogalactan Sugar Chain Induces Pollen-Tube Competency to Respond to Ovular Guidance. *Current Biology*, *26*(8), 1091–1097. https://doi.org/10.1016/j.cub.2016.02.040

Mizuta, Y., & Higashiyama, T. (2018). Chemical signaling for pollen tube guidance at a glance. *Journal of Cell Science*, *131*(2). https://doi.org/10.1242/jcs.208447

Noyszewski, A. K., Liu, Y.-C., Tamura, K., & Smith, A. G. (2017). Polymorphism and structure of style–specific arabinogalactan proteins as determinants of pollen tube growth in Nicotiana. *BMC Evolutionary Biology*, *17*(1), 186. https://doi.org/10.1186/s12862-017-1011-2

Ogawa-Ohnishi, M. & Matsubayashi, Y. (2015). Identification of three potent hydroxyproline O-galactosyltransferases in Arabidopsis. *Plant J.* 81(5):736-46. doi: 10.1111/tpj.12764. PMID: 25600942.

Okuda, S., Tsutsui, H., Shiina, K., Sprunck, S., Takeuchi, H., Yui, R., Kasahara, R. D., Hamamura, Y., Mizukami, A., Susaki, D., Kawano, N., Sakakibara, T., Namiki, S., Itoh, K., Otsuka, K., Matsuzaki, M., Nozaki, H., Kuroiwa, T., Nakano, A., … Higashiyama, T. (2009). Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells. *Nature*, *458*(7236), 357–361. https://doi.org/10.1038/nature07882

Palanivelu, R., & Tsukamoto, T. (2012). Pathfinding in angiosperm reproduction: Pollen tube guidance by pistils ensures successful double fertilization. *Wiley Interdisciplinary Reviews. Developmental Biology*, *1*(1), 96–113. https://doi.org/10.1002/wdev.6

Pasternak, T., Tietz, O., Rapp, K., Begheldo, M., Nitschke, R., Ruperti, B., & Palme, K. (2015). Protocol: An improved and universal procedure for whole-mount immunolocalization in plants. *Plant Methods*, *11*(1), 50. https://doi.org/10.1186/s13007-015-0094-2

Pattison, R. J., Csukasi, F., Zheng, Y., Fei, Z., Knaap, E. van der, & Catalá, C. (2015). Comprehensive Tissue-Specific Transcriptome Analysis Reveals Distinct Regulatory Programs during Early Tomato Fruit Development. *Plant Physiology*, *168*(4), 1684–1701. https://doi.org/10.1104/pp.15.00287

Pennell, R. I., Janniche, L., Kjellbom, P., Scofield, G. N., Peart, J. M., & Roberts, K. (1991). Developmental Regulation of a Plasma Membrane Arabinogalactan Protein Epitope in Oilseed Rape Flowers. *The Plant Cell*, *3*(12), 1317–1326. https://doi.org/10.1105/tpc.3.12.1317

Pereira, Ana M., Lopes, A. L., & Coimbra, S. (2016). Arabinogalactan Proteins as Interactors along the Crosstalk between the Pollen Tube and the Female Tissues. *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.01895

Pereira, Ana Marta, Masiero, S., Nobre, M. S., Costa, M. L., Solís, M.-T., Testillano, P. S., Sprunck, S., & Coimbra, S. (2014). Differential expression patterns of arabinogalactan proteins in Arabidopsis thaliana reproductive tissues. *Journal of Experimental Botany*, *65*(18), 5459–5471. https://doi.org/10.1093/jxb/eru300

Pitto, L., Giorgetti, L., Turrini, A., Evangelista, M., Luccarini, G., Colella, C., Collina, F., Caltavuturo, L., & Nuti Ronchi, V. (2001). Floral genes expressed in tomato hypocotyl explants in liquid culture. *Protoplasma*, *218*(3), 168–179. https://doi.org/10.1007/BF01306606

Reimann, R., Kah, D., Mark, C., Dettmer, J., Reimann, T. M., Gerum, R. C., Geitmann, A., Fabry, B., Dietrich, P., & Kost, B. (2020). Durotropic Growth of Pollen Tubes. *Plant Physiology*, *183*(2), 558–569. https://doi.org/10.1104/pp.19.01505

Rydahl, M. G., Hansen, A. R., Kračun, S. K., & Mravec, J. (2018). Report on the Current Inventory of the Toolbox for Plant Cell Wall Analysis: Proteinaceous and Small Molecular Probes. *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.00581

Sato, S., Tabata, S., Hirakawa, H., Asamizu, E., Shirasawa, K., Isobe, S., Kaneko, T., Nakamura, Y., Shibata, D., Aoki, K., Egholm, M., Knight, J., Bogden, R., Li, C., Shuang, Y., Xu, X., Pan, S., Cheng, S., Liu, X., … Universitat Pompeu Fabra. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, *485*(7400), 635–641. https://doi.org/10.1038/nature11119

Schultz, C., Gilson, P., Oxley, D., Youl, J., & Bacic, A. (1998). GPI-anchors on arabinogalactan-proteins: Implications for signalling in plants. *Trends in Plant Science*, *3*(11), 426–431. https://doi.org/10.1016/S1360-1385(98)01328-4

Schultz, C. J., Rumsewicz, M. P., Johnson, K. L., Jones, B. J., Gaspar, Y. M., & Bacic, A. (2002). Using Genomic Resources to Guide Research Directions. The Arabinogalactan Protein Gene Family as a Test Case. *Plant Physiology*, *129*(4), 1448–1463. https://doi.org/10.1104/pp.003459

Seifert, G. J. (2018). Fascinating Fasciclins: A Surprisingly Widespread Family of Proteins that Mediate Interactions between the Cell Exterior and the Cell Surface. *International Journal of Molecular Sciences*, *19*(6). https://doi.org/10.3390/ijms19061628

Selinski, J., & Scheibe, R. (2014). Pollen tube growth: Where does the energy come from? *Plant Signaling & Behavior*, *9*(12), e977200. https://doi.org/10.4161/15592324.2014.977200

Shen, S., Ma, S., Liu, Y., Liao, S., Li, J., Wu, L., Kartika, D., Mock, H.-P., & Ruan, Y.-L. (2019). Cell Wall Invertase and Sugar Transporters Are Differentially Activated in Tomato Styles and Ovaries During Pollination and Fertilization. *Frontiers in Plant Science*, *10*. https://doi.org/10.3389/fpls.2019.00506

Showalter, A. M., Gao, M., Kieliszewski, M. J., & Lamport, D. T. A. (2000). Characterization and Localization of a Novel Tomato Arabinogalactan-Protein (LeAGP-1) and the Involvement of Arabinogalactan-Proteins in Programmed Cell Death. In E. A. Nothnagel, A. Bacic, & A. E. Clarke (Eds.), *Cell and Developmental Biology of Arabinogalactan-Proteins* (pp. 61–70). Springer US. https://doi.org/10.1007/978-1-4615-4207-0_6

Showalter, A. M., Keppler, B., Lichtenberg, J., Gu, D., & Welch, L. R. (2010). A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. *Plant Physiology*, *153*(2), 485–513. https://doi.org/10.1104/pp.110.156554

Smith, A. G., Eberle, C. A., Moss, N. G., Anderson, N. O., Clasen, B. M., & Hegeman, A. D. (2013). The transmitting tissue of Nicotiana tabacum is not essential to pollen tube growth, and its

ablation can reverse prezygotic interspecific barriers. *Plant Reproduction*, *26*(4), 339–350. https://doi.org/10.1007/s00497-013-0233-8

Sprunck, S., Hackenberg, T., Englhart, M., & Vogler, F. (2014). Same same but different: Sperm-activating EC1 and ECA1 gametogenesis-related family proteins. *Biochemical Society Transactions*, *42*(2), 401–407. https://doi.org/10.1042/BST20140039

Sprunck, S., Rademacher, S., Vogler, F., Gheyselinck, J., Grossniklaus, U., & Dresselhaus, T. (2012). Egg Cell–Secreted EC1 Triggers Sperm Cell Activation During Double Fertilization. *Science*, *338*(6110), 1093–1097. https://doi.org/10.1126/science.1223944

Su, S., & Higashiyama, T. (2018). Arabinogalactan proteins and their sugar chains: Functions in plant reproduction, research methods, and biosynthesis. *Plant Reproduction*, *31*(1), 67–75. https://doi.org/10.1007/s00497-018-0329-2

Sun, W., Kieliszewski, M. J., & Showalter, A. M. (2004). Overexpression of tomato LeAGP-1 arabinogalactan-protein promotes lateral branching and hampers reproductive development. *The Plant Journal*, *40*(6), 870–881. https://doi.org/10.1111/j.1365-313X.2004.02274.x

Takeuchi, H., & Higashiyama, T. (2011). Attraction of tip-growing pollen tubes by the female gametophyte. *Current Opinion in Plant Biology*, *14*(5), 614–621. https://doi.org/10.1016/j.pbi.2011.07.010

Takeuchi, H., & Higashiyama, T. (2012). A Species-Specific Cluster of Defensin-Like Genes Encodes Diffusible Pollen Tube Attractants in Arabidopsis. *PLoS Biology*, *10*(12). https://doi.org/10.1371/journal.pbio.1001449

Twarda-Clapa, A., Labuzek, B., Krzemien, D., Musielak, B., Grudnik, P., Dubin, G., & Holak, T. A. (2018). Crystal structure of the FAS1 domain of the hyaluronic acid receptor stabilin-2. *Acta Crystallographica Section D: Structural Biology*, *74*(7), 695–701. https://doi.org/10.1107/S2059798318007271

Vogler, F., Schmalzl, C., Englhart, M., Bircheneder, M., & Sprunck, S. (2014). Brassinosteroids promote Arabidopsis pollen germination and growth. *Plant Reproduction*, *27*(3), 153–167. https://doi.org/10.1007/s00497-014-0247-x

Wei, L. Q., Xu, W. Y., Deng, Z. Y., Su, Z., Xue, Y., & Wang, T. (2010). Genome-scale analysis and comparison of gene expression profiles in developing and germinated pollen in Oryza sativa. *BMC Genomics*, *11*(1), 338. https://doi.org/10.1186/1471-2164-11-338

Wingett, S. W., & Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research*, *7*, 1338. https://doi.org/10.12688/f1000research.15931.2

Wolters-Arts, M., Lush, W. M., & Mariani, C. (1998). Lipids are required for directional pollen-tube growth. *Nature*, *392*(6678), 818–821. https://doi.org/10.1038/33929

Wu, H.-M., Wang, H., & Cheung, A. Y. (1995). A pollen tube growth stimulatory glycoprotein is deglycosylated by pollen tubes and displays a glycosylation gradient in the flower. *Cell*, *82*(3), 395–403. https://doi.org/10.1016/0092-8674(95)90428-X

Wu, Y., Fan, W., Li, X., Chen, H., Takáč, T., Šamajová, O., Fabrice, M. R., Xie, L., Ma, J., Šamaj, J., & Xu, C. (2017). Expression and distribution of extensins and AGPs in susceptible and resistant

banana cultivars in response to wounding and Fusarium oxysporum. *Scientific Reports*, *7*(1), 42400. https://doi.org/10.1038/srep42400

Yariv, J., Rapport, M., & Graf, L. (1962). The interaction of glycosides and saccharides with antibody to the corresponding phenylazo glycosides. *Biochemical Journal*, *85*(2), 383–388. https://doi.org/10.1042/bj0850383

Yates, E. A., Valdor, J.-F., Haslam, S. M., Morris, H. R., Dell, A., Mackie, W., & Knox, J. P. (1996). Characterization of carbohydrate structural features recognized by anti-arabinogalactan-protein monoclonal antibodies. *Glycobiology*, *6*(2), 131–139. https://doi.org/10.1093/glycob/6.2.131

Zhang, M. J., Zhang, X. S., & Gao, X.-Q. (2020). ROS in the Male–Female Interactions During Pollination: Function and Regulation. *Frontiers in Plant Science*, *11*. https://doi.org/10.3389/fpls.2020.00177

Zhang, S., Xu, M., Qiu, Z., Wang, K., Du, Y., Gu, L., & Cui, X. (2016). Spatiotemporal transcriptome provides insights into early fruit development of tomato (Solanum lycopersicum). *Scientific Reports*, 6, 23173. https://doi.org/10.1038/srep23173

Zheng, R. H., Su, S. D., Xiao, H., & Tian, H. Q. (2019). Calcium: A Critical Factor in Pollen Germination and Tube Elongation. *International Journal of Molecular Sciences*, *20*(2). https://doi.org/10.3390/ijms20020420

Zhong, Y., & Shanley, J. (1995). Altered nerve terminal arborization and synaptic transmission in Drosophila mutants of cell adhesion molecule fasciclin I. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *15*(10), 6679–6687.

Figure legends

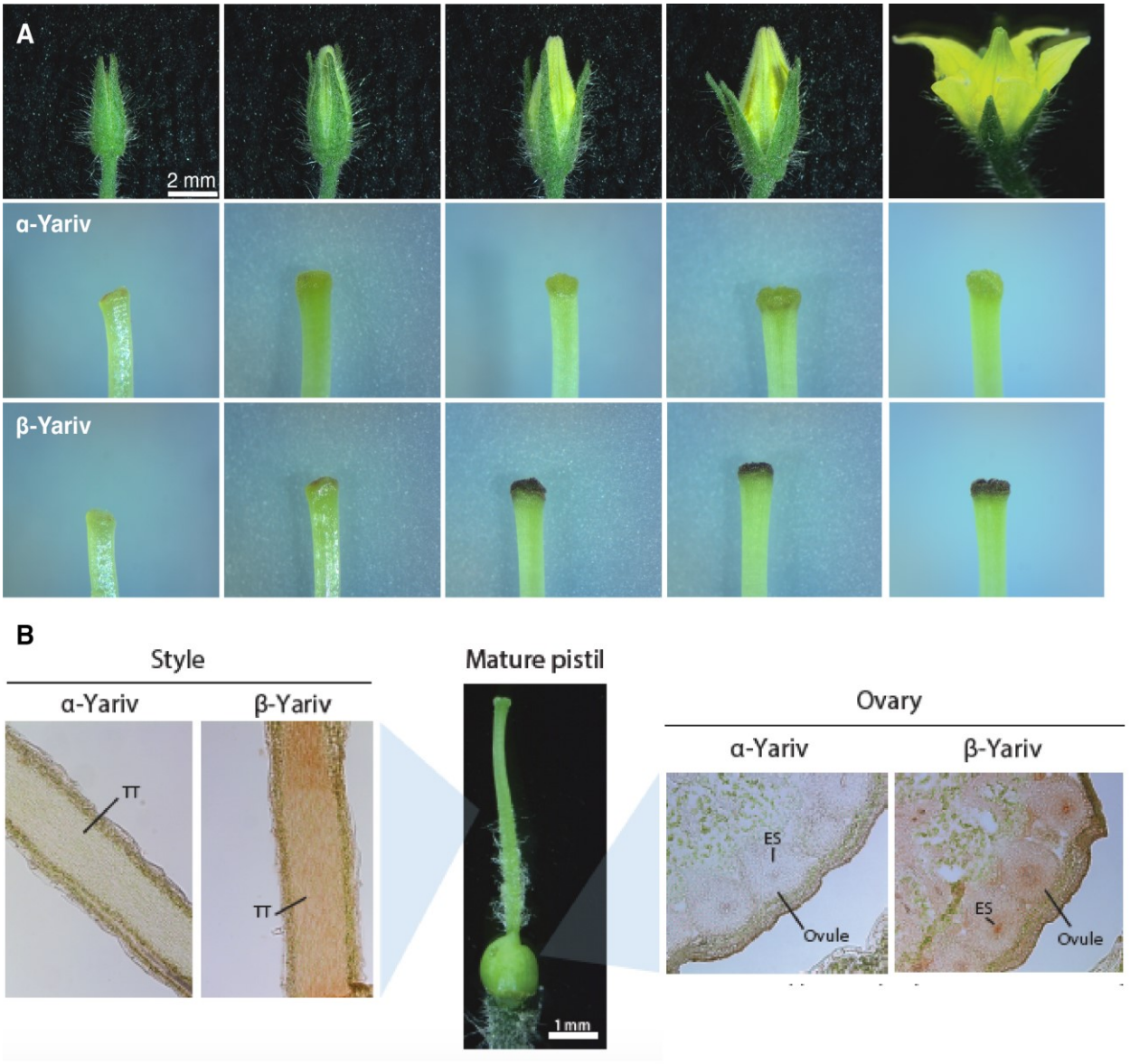


Figure 1. Arabinogalactan glycoproteins accumulate in the mature tomato pistil. A) Pistil in different developmental stages (top) were dissected and stained with α (middle) or β-Yariv (bottom). AGPs react with the β-Yariv reagent, visible by the accumulation of a brown-red precipitate. α-Yariv stained samples serve as negative control, as α-Yariv does not react with AGPs. **B)** In cryosections of mature pistils, β-Yariv reactive AGPs accumulated in the transmitting tract in the style (longitudinal section) and in the ovary (cross-section), visible staining is present in the embryo sac and placenta.

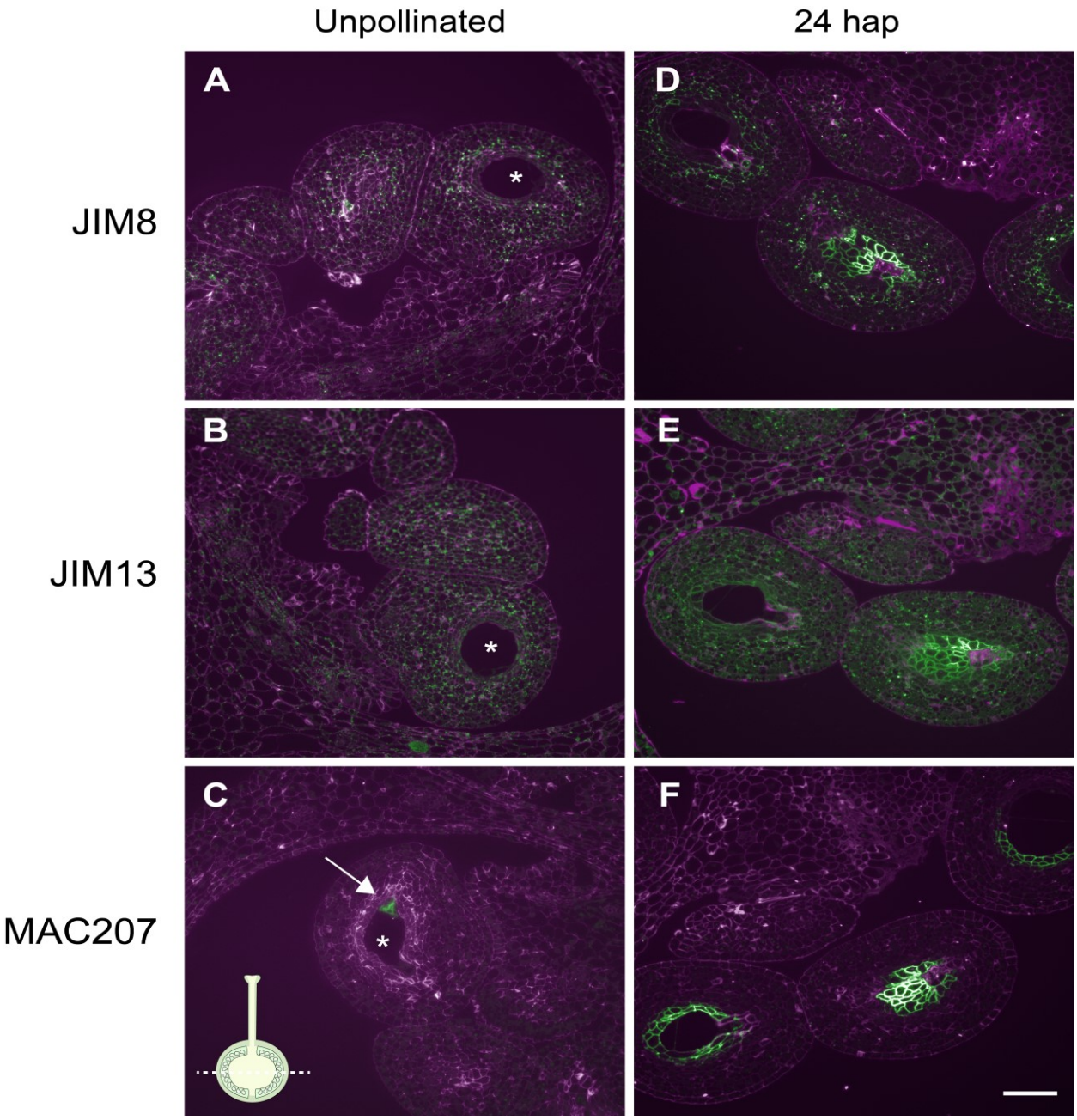


Figure 2. Arabinogalactan glycoproteins in the ovule display a distinctive pattern pre- and post-fertilization. Representative micrographs of semi-thin sections labeled with anti-AGP glycan antibodies (green) and calcofluor white as a counterstain for the cell wall (magenta). Pictures A to C correspond to samples sectioned from unpollinated pistils, from D to F, samples sectioned from manually pollinated pistils, collected 24 h after pollination. The arrow indicates the signal detected when labeling with MAC207; Stars mark the embryo sac. Scale bar 50 μm. Pistil schematic on C represents section orientation. Note that the embryo sac is difficult to capture in sectioned material due to its large size.

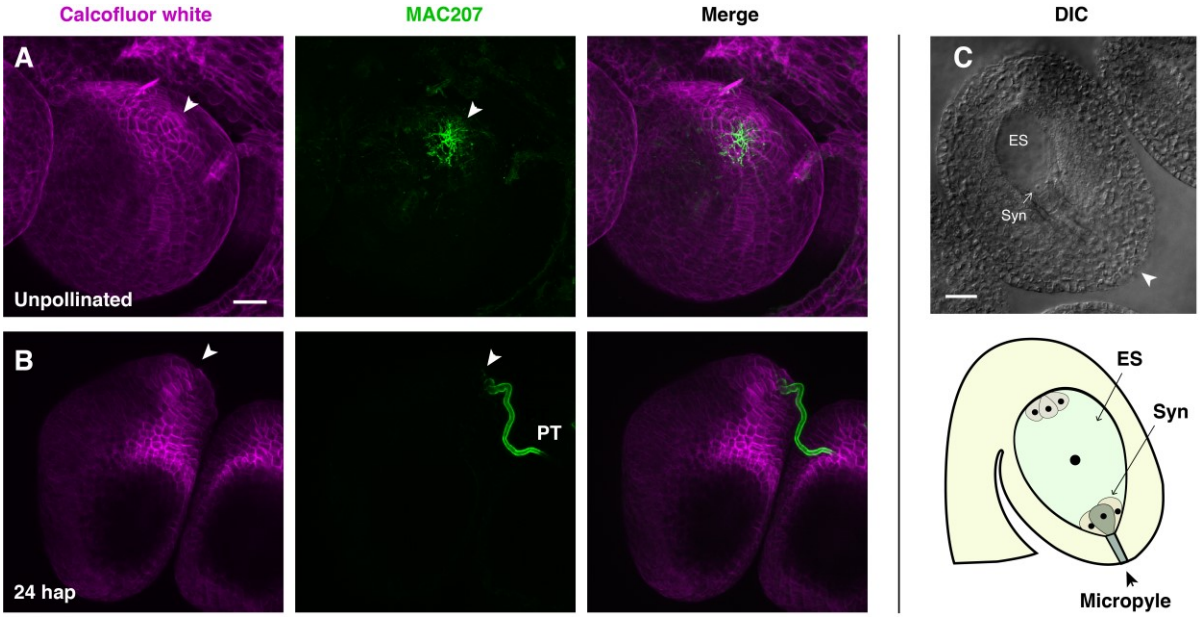


Figure 3. Arabinogalactan glycoproteins in the micropyle surface is not detectable post-fertilization. Confocal laser microscopy of whole mount immunofluorescence of ovules from unpollinated pistils (A) or 24 h after pollination (B, 24 hap). MAC207 reactive AGPs accumulated in the micropyle (arrowhead) in ovules from unpollinated pistils. After a pollen tube (PT) had fertilized the ovule, the diffuse micropylar AGP signal was no longer detectable. MAC207 also bound to AGPs present in the pollen tube. Magenta channel corresponds to calcofluor white, staining the cell wall for contrast. (C) DIC micrograph of cleared ovules to better appreciate the ovule anatomy; bottom, schematic of the tomato ovule. ES: Embryo sac, Syn: synergids.

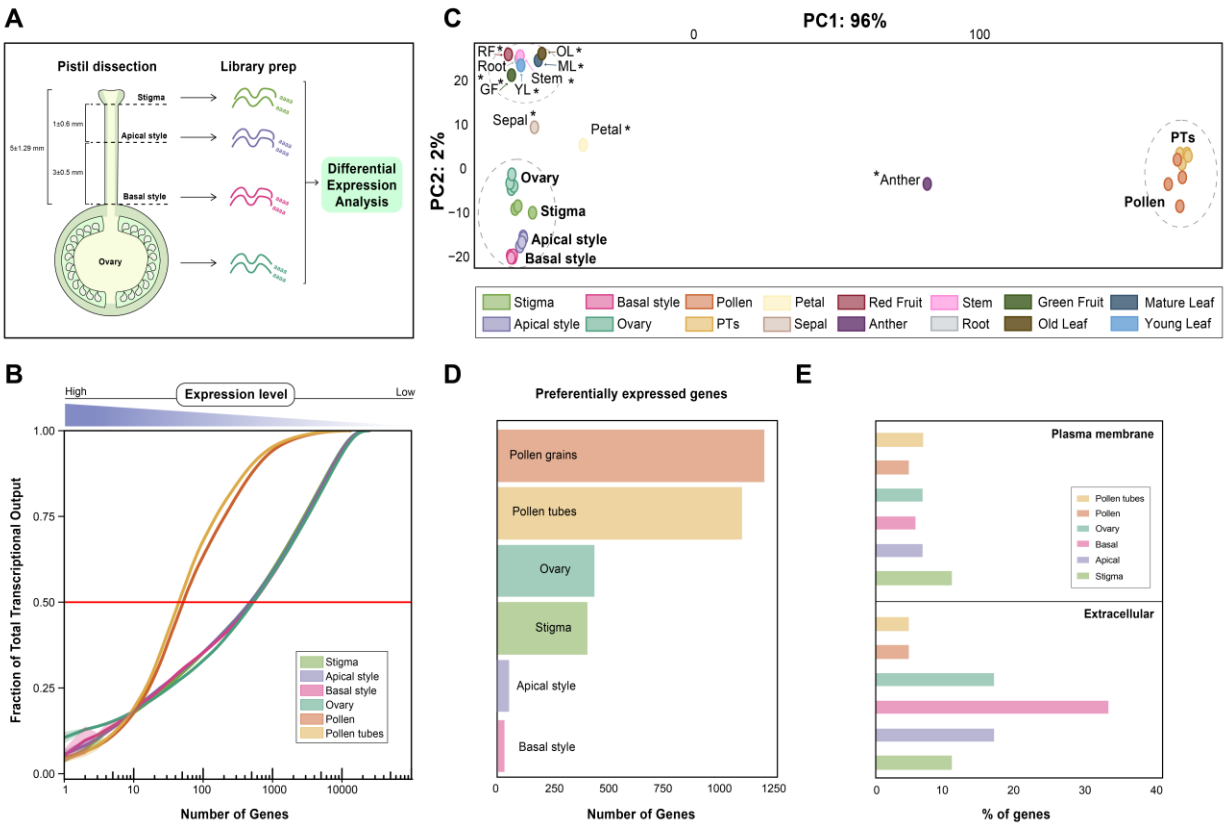


Figure 4. Exploratory analysis of transcriptomes from reproductive tissue of tomato. A) Pipeline for library generation showing subsections of the tomato pistil used to generate libraries. Pistils of pre-anthesis flowers were emasculated, covered and allowed to mature for 24 h before dissection into the four subsections comprising the stigma, the apical style, the basal style and the ovary. B) Transcriptome complexity of tomato reproductive tissues. Cumulative distribution of the mean fraction of total transcription per tissue, contributed by genes sorted from highest expressed to lowest expressed (X axis, log10 scale). Per tissue, the lines represent the mean expression values across biological replicates, divided by the sum of all mean expression values from the tissue (Fraction of Total Transcriptional Output, Y axis). The shaded area surrounding each line represents the dispersion of the data, calculated by dividing the standard deviation by the cumulative sum of all mean expression values. The red line indicates half of the Total Transcriptional Output to facilitate comparison. C) Principal Component Analysis (PCA) of generated datasets (bold letters) and published transcriptomes form vegetative tissues and other floral parts (stars). Major clusters in gray dashed circles. Published transcriptomes taken from (Ezura *et al.*, 2017). RF: Red Fruit, GF: Green Fruit, YL: Young Leaf, ML: Mature Leaf, OL: Old Leaf, PTs: Pollen tubes in the tri-cellular stage. D) Preferentially expressed genes identified by differential gene expression analysis. Each sample

was compared to the rest of the tissues and non-overlapping genes were considered as preferentially expressed. E) Percentage of preferentially expressed genes per sample with predicted plasma membrane or extracellular (secreted) localization.

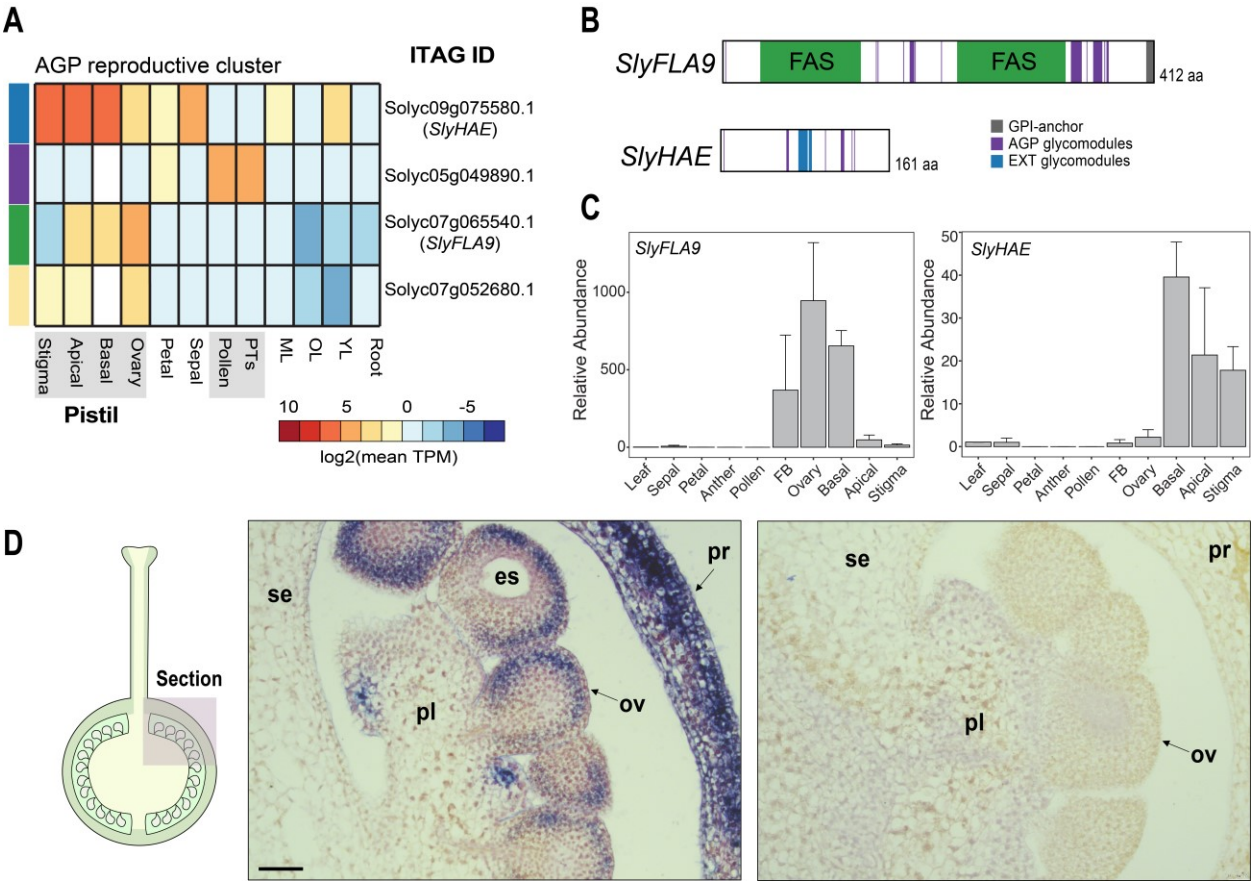
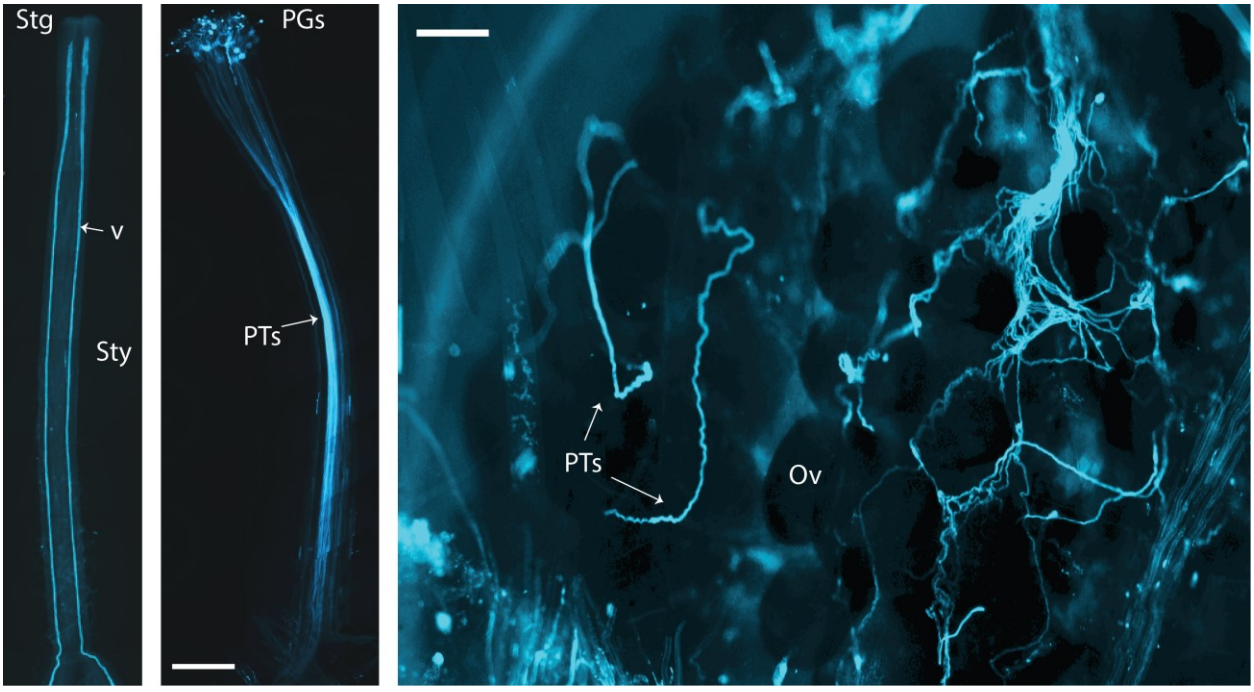
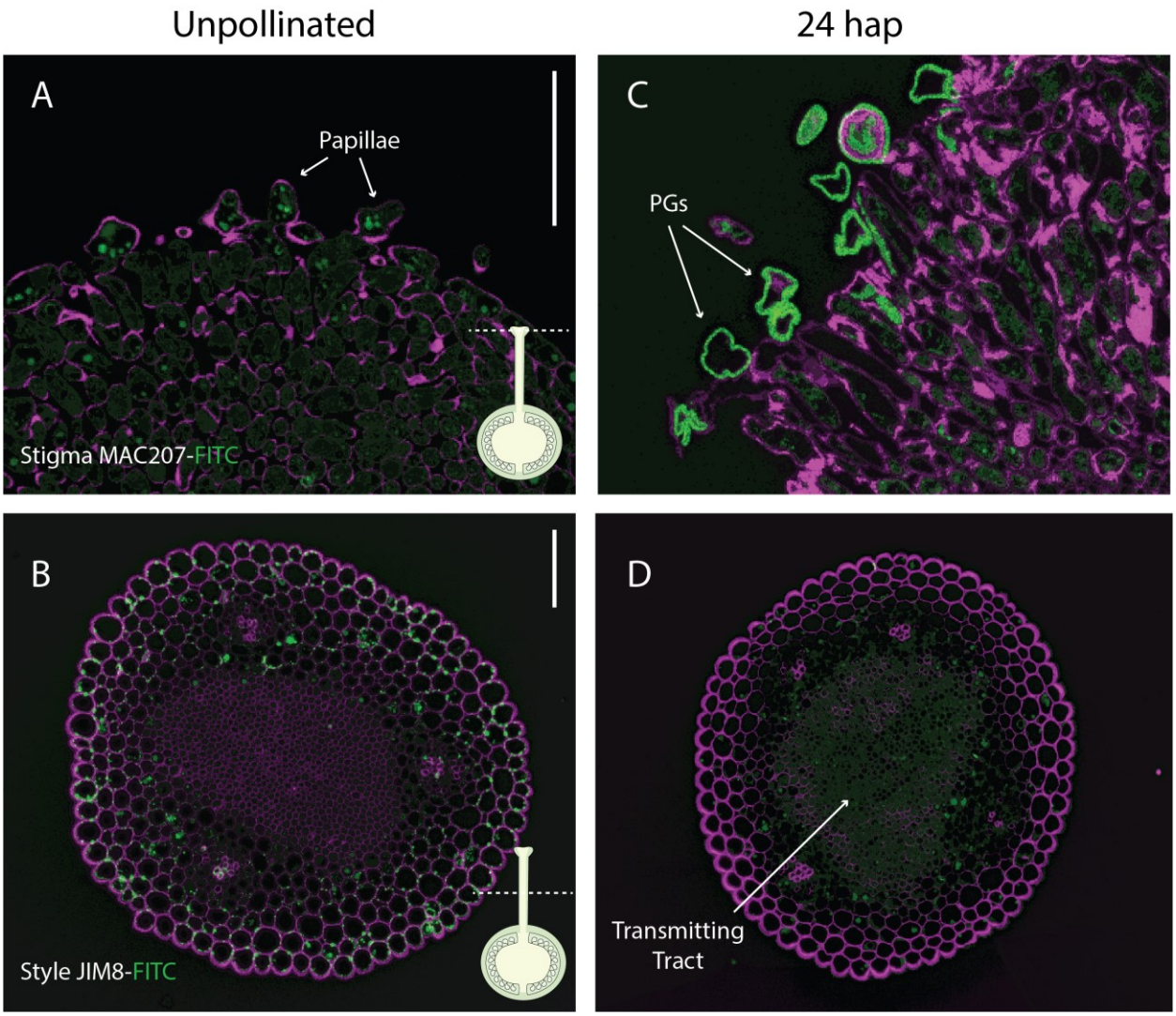


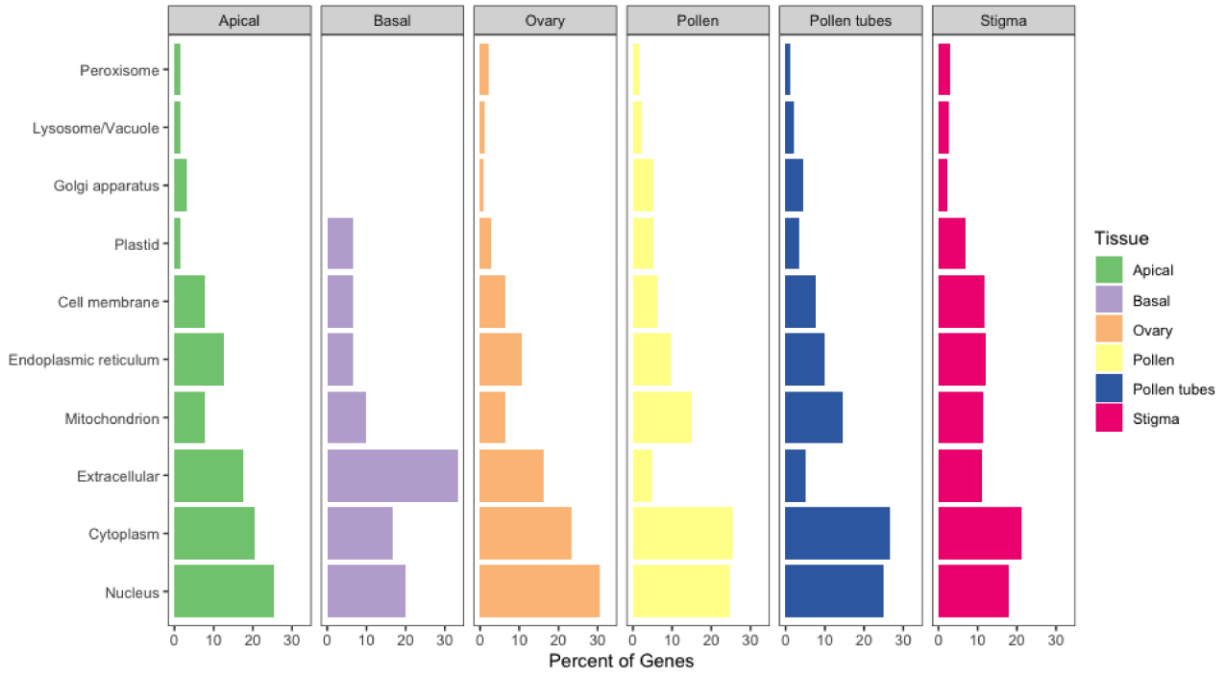
Figure 5. Expression profiles of pistil AGPs in tomato. A) Expression profiles of pistil AGPs represented in our datasets. The heatmap corresponds to a subcluster taken from Sup. Fig. 4, where AGP genes had higher expression in reproductive tissues compared to the rest of the tissues. Color code on the left represents distinct classes of AGPs. Blue: AGP-Extensin hybrid; purple: classical AGP; green: Fasciclin-like AGP; yellow: Lysine-rich AGP. **B)** Protein architecture of Fasciclin-like AGP (*SlyFLA9*, *Solyc07g065540.1)* and hybrid AGP-Extensin (*SlyHAE*, *Solyc09g075580.1*). **C)** Quantitative RT-PCR of *SlyFLA9* and *SlyHAE* across tomato tissues. Expression values were normalized using *SlyACT* (*Solyc00g017210.1*) and *SlyUBI* (*Solyc01g056940.2*) as references. FB: 2 mm flower buds. **D)** *In situ* detection of the *SlyFLA9* transcript in a sagittal section of the ovary, dissected from an unpollinated pistil. Left: schematic of the pistil sections shown for reference, middle: antisense probe, right: sense probe; es: embryo sac, se: septum, pl: placenta, ov: ovule, pr: pericarp. Scale bar 50 µm.



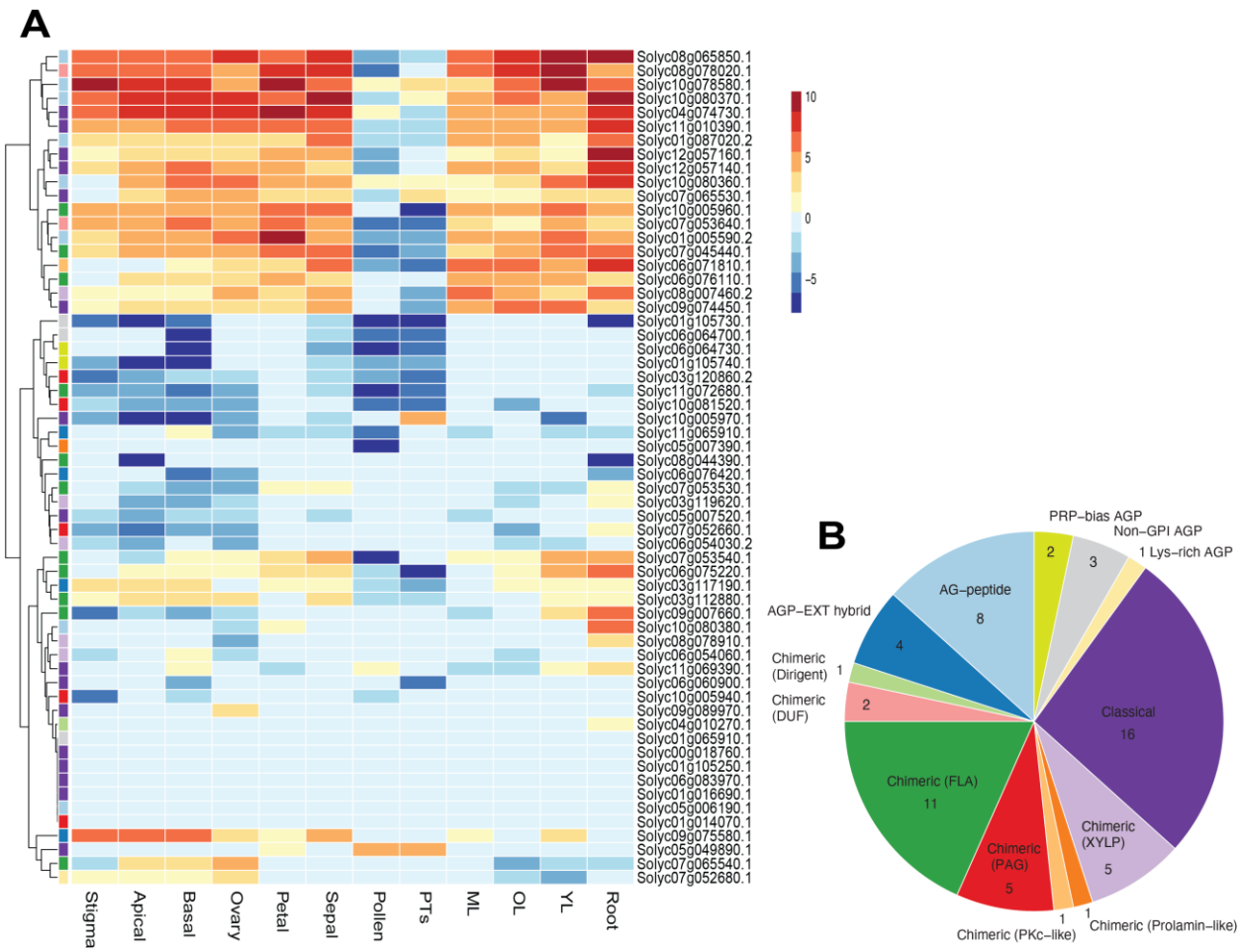
Supplemental Figure 1. Pollen tubes reach ovules and fertilize them after 24 h post manual pollination. Unpollinated pistils (left) or 24 h after pollination (middle) were harvested, fixed and stained with aniline blue fluorochrome to visualize pollen tubes growing through pistil tissues. In unpollinated pistils, vascular bundles were visible post-staining. Intact pollen tubes targeting ovules were visualized by dissecting the pericarp of 24 hap pistils (right). Stg: stigma, Sty: style, v: vasculature, PGs: pollen grains, PTs: pollen tubes, Ov: ovule. Scale bar in middle and right panel corresponds to 500 µm.



Supplemental Figure 2. Representative images of immunolocalization of AGPs in the stigma and style during pollen pistil-interaction. In the style, punctate signal in the papillae was observed in unpollinated pistils when staining with MAC207 (A). Punctate signal in the papillae was not observed post pollination, whereas strong signal was detected in pollen grains adhered to them (B-D). In the style, punctate signal in the cells surrounding the secretory cells of the transmitting tract were observed when staining with JIM8 in unpollinated pistils (B). After pollination, only weak signal was detected (F-H). Scale bar in A is 20 μm; scale bar in E is 50 μm.



Supplemental Figure 3. Predicted subcellular localization of preferentially expressed genes from pistil subsection and pollen grains/tubes. Subcellular localization was predicted using DeepLoc-01 (Almagro Armenteros et al., 2017).



Supplemental Figure 4. A) Hierarchical clustering of AGPs gene expression across tomato tissues. Colors next to the dendrogram, represent the classification color coded as in B. The heatmap represents the expression patterns observed for the 60 putative AGP genes across tomato tissues. The color scale in the heatmap represents the log2 mean TPM value per gene. B) Classification of the AGP genes in A. The majority of genes identified belong to the classical AGP group, followed by Fasciclin-like AGPs (FLA). PAG: phytocyanin-like AGPs; DUF: Domain of Unknown Function AGPs; XYLP: Xylogen-like Protein AGPs; PKc: Protein Kinase-like AGPs; PRP-bias: Proline-rich protein bias AGP; AGP-EXT: AGP Extensin hybrid. C) Protein architecture of the reproductive AGP cluster which includes Solyc09g075580.1, Solyc07g065540.1 and Solyc07g052680.1 expressed in the pistil and Solyc05g049890.1, expressed in pollen grains and tubes.

>Solyc09g075580.1 SlyHAE

MFAPNSQFQFFLLLFFSFSLNNVQITQAMSTITAISKDQIACEMCTTCENPCQPIFPPPPPPPSPSPPLLCPPPPSPPP
PPPPSPPPLPPVNYCPPPPPPPARRSCPEDCSLQPRSPPYSIYPYFSPPSTYPNKSTRYKNQPIVTYLIIAIACLFKTF
GLM*

>Solyc07g065540.1 SlyFLA9

MQLPSSATVAAVVLSVVLFICATEAHNITHILADHKQFSTFNHYLTTTHLAAEINRRQTITVCAVDNAGMSDLLSKQLSI
YTIKNVLSFHVLLDYFDAKKLHQITNGTALAATMFQATGSATGSSGFVNITDLRGGKVGLSPADYNGPPPAKFVKSIAEI
PYNISVIQISTILPSDEAEAPTPGPSQMNLTSLMSAKGCKVFAETLLASPAEKTFEDNVDGGVTIFCPRDDAMKKFLPKF
KNLTAEGKQSLLEYHGIPIYQSISNLKSNNGDMNTLATDGAKKYAVVIQNDGEDVTIKTKIVTAKITATVVDKLPLAIYS
LDKVLLPEELFKASPTPAPAPAPEARAESPKHSKSPPAPASPAESPADSPADGPNGDADDLTDSGAVKYNAGASLAAVFS
LWFAFNVLMEV*

>Solyc07g052680.1 Lys-rich-AGP

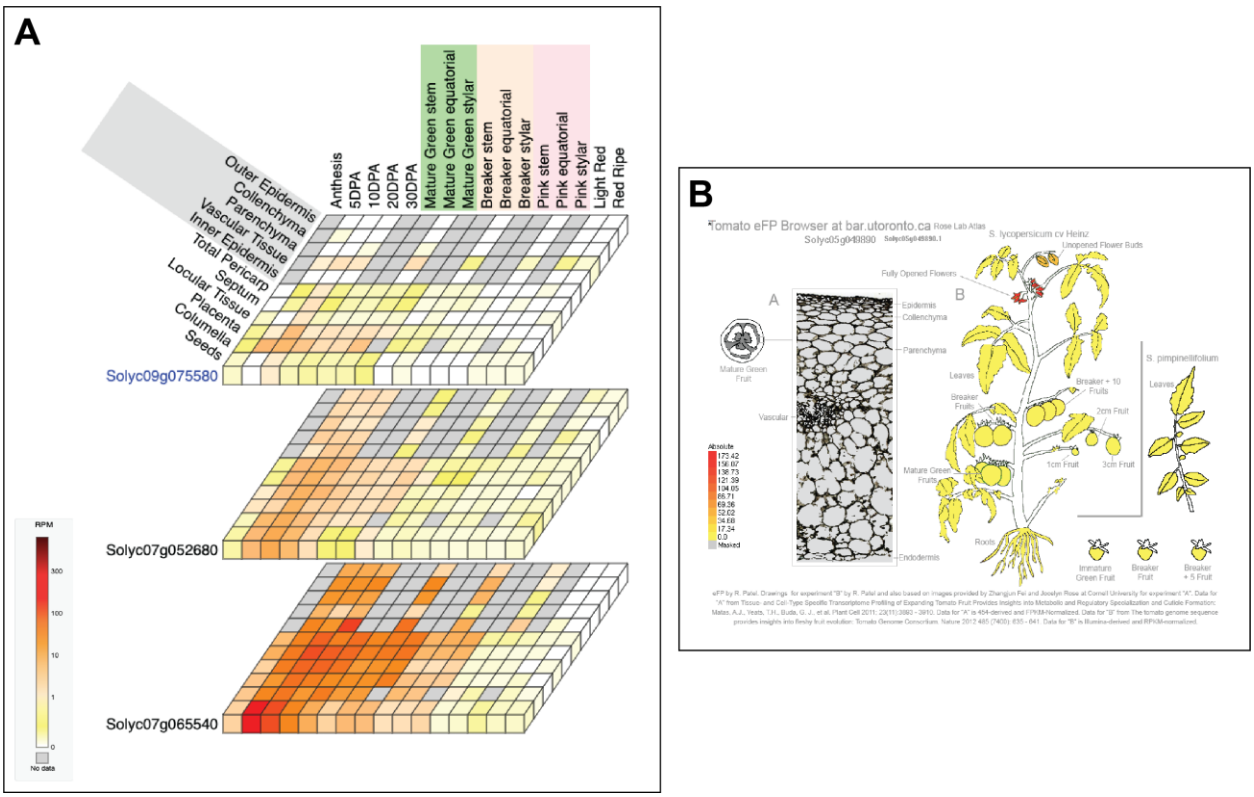
MCYVGKATKIFIFIVTVLVVTGLILLGFGLLRHHNQKGENKCSGDSCDQNQYQSPIVYPPPTTSTNTNNPISPLPISTPSQ
PNTPNPNLPQPPPPPSPDNPTPETPNLTPPPPPDTVVSTPPPLPPPAVSLTPPPTLSPPSPVTVSPGPVQS*

>Solyc05g049890.1 Pollen classical AGP

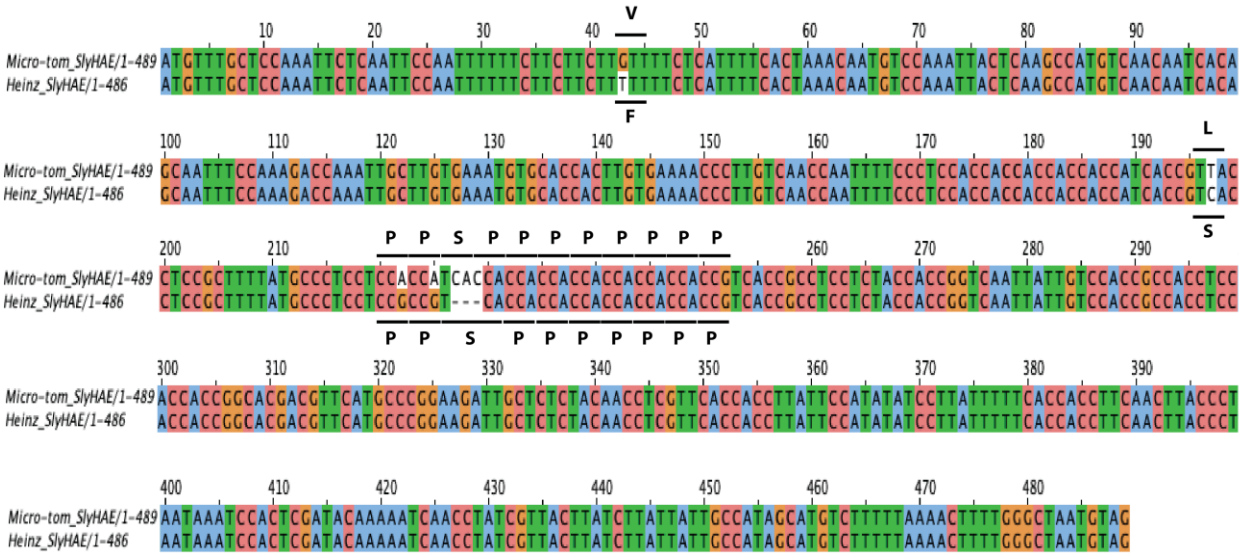
MKIIITPFFLLVLLLSRMETSEAAKLPPARSLAPHPAVSPSPRPAGVSPSPHPAEKNSVGLKKVSSAPAPLPPTTKKNNVP
LVSNKSKTRA*

XP AGP glycomodules (X= Ser, Ala, Thr, Gly, Val)
SP (3-5) Extensin glycomodules
FAS Fasciclin-like domain
GPI anchor
L Lysine residue

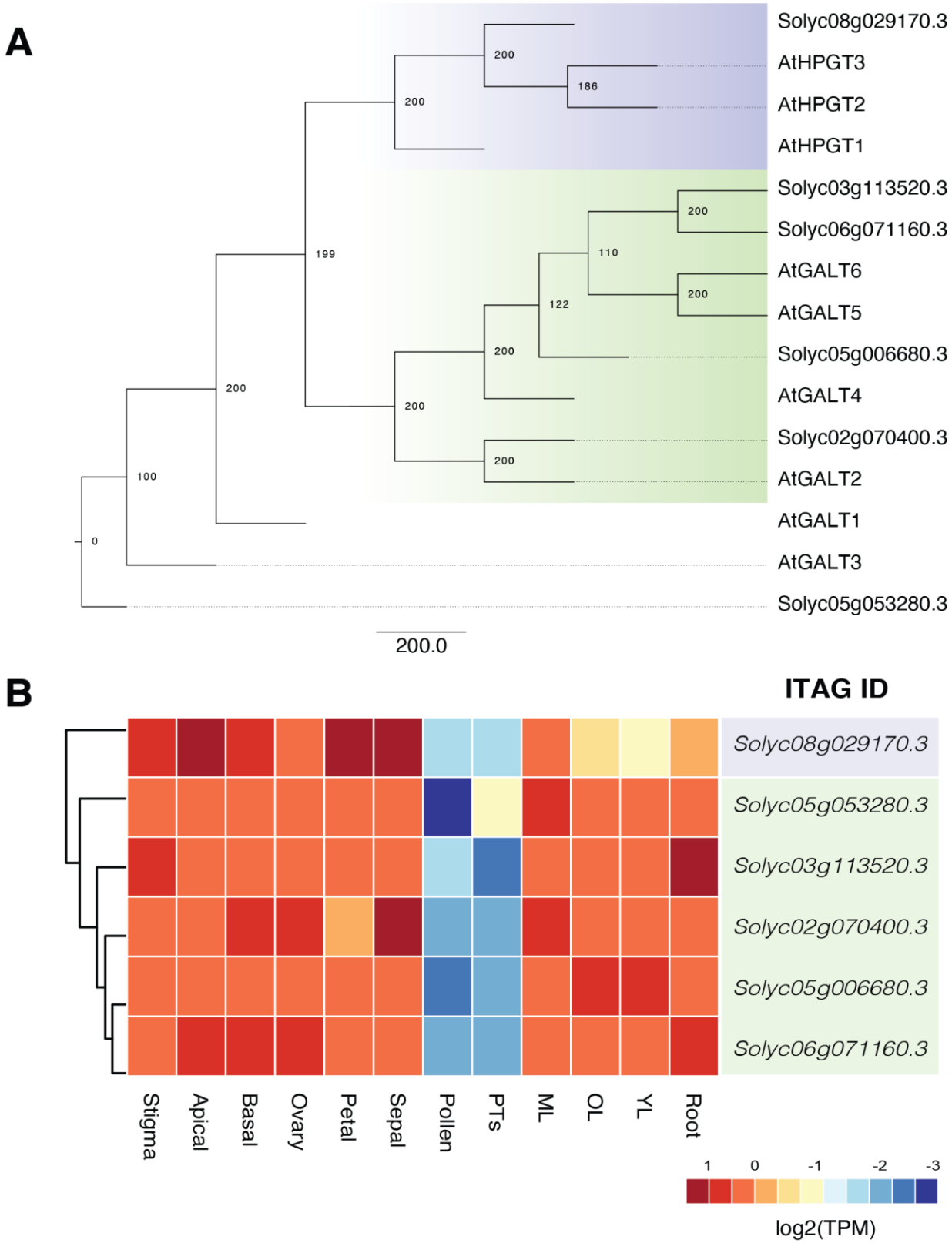
Supplemental Figure 5. Predicted protein sequences of the reproductively expressed AGPs with key regions marked.



Supplemental Figure 6. A) Expression of pistil AGPs (from Transcriptome Expression Atlas -- TEA from SolGenomics Network) and B) Pollen and pollen tube expressed AGP (from Tomato eFP browser).

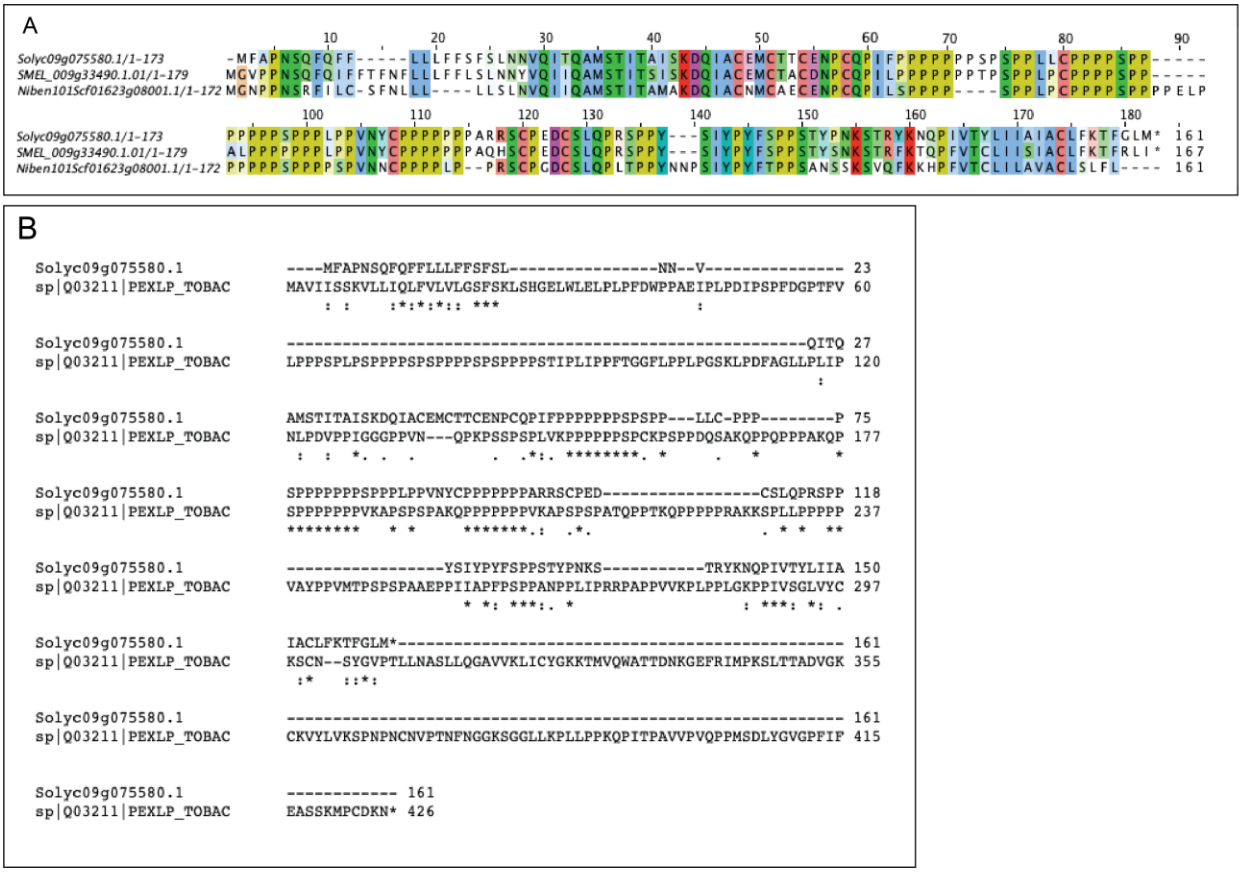


Supplemental Figure 7. *SlyHAE* (Solyc09g07558.1) reported polymorphisms in the coding region between the Micro-tom (top) and Heinz 1706 (bottom) cultivars. Indicated in bold letters are the corresponding amino acids for each sequence., The polymorphisms in the Micro-tom *SlyHAE* sequence were confirmed by Sanger sequencing.



Supplemental Figure 8. A) Neighbor joining phylogenetic tree of hydroxyproline

galactosyltransferases (hyp-GALT and HPGT) proteins from Arabidopsis (At) and tomato (Solyc). Values in each node represent bootstrap values. B) Gene expression of tomato GALT/HPGTs across tissues. HPGTs and GALTs are highlighted in blue and green respectively. The color scale in the heatmap represents the log2 mean TPM value per gene.



Supplemental Figure 9. A) Alignment of the *SlyHAE* (Solyc09g07558.1) with homologs in eggplant (*SMEL_009g33490.1.01*) and *N. benthamiana* (*Niben101Scf01623g08001.1*). B) Protein sequence alignment between *SlyHAE* and tobacco PELPIII (PEXLP_TOBAC).