Differential expression patterns of Arabinogalactan Proteins in Arabidopsis thaliana reproductive tissues

Running title: AGPs in Arabidopsis reproductive tissues

Ana Marta Pereira¹, ², ³, Simona Masiero³, Margarida Sofia Nobre¹, Mário Luís Costa¹, ², María-Teresa Solís⁴, Pilar S. Testillano⁴, Stefanie Sprunck⁵ and Sílvia Coimbra¹, ², ⁶

¹Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre, 4169-007 Porto, Portugal,
²Center for Biodiversity, Functional & Integrative Genomics (BioFIG), http://biofig.fc.ul.pt
³Dipartimento di BioScienze, Università degli Studi di Milano, 20133 Milan, Italy
⁴Pollen Biotechnology of Crop Plants Group, Centro de Investigaciones Biológicas (CIB) CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain
⁵Cell Biology and Plant Biochemistry, Biochemie-Zentrum Regensburg, University of Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany
⁶To whom correspondence should be addressed: Fax: +351 220402799. E-mail: scoimbra@fc.up.pt

Date of submission: 07/04/2014

Total word count: 9433 (including main body text, references and figure legends)
Total number of figures: 6 (all pictures to be print in color)
Supplemental material: 2 tables
Arabinogalactan proteins (AGPs) are heavily glycosylated proteins existing in all plant kingdom and differentially distributed through distinctive developmental stages. Here we show the individual distribution of specific *Arabidopsis* AGPs: AGP1, AGP9, AGP12, AGP15 and AGP23, throughout reproductive tissues and point out their possible roles in several reproductive processes. AGP genes specifically expressed in the female tissues were identified using available microarray data. This selection was confirmed by promoter analysis using multiple GFP fusions to a nuclear localization signal (NLS), GUS fusions, and *in situ* hybridization as an approach to confirm the AGPs expression patterns. Promoter analysis allowed the detection of a specific and differential presence of these proteins along the pathway followed by the pollen tube during its journey to reach the egg and the central cell inside the embryo sac. AGP1 is expressed in the stigma, the style, the transmitting tract, the funiculus, and in the chalazal and funiculus tissues of the ovules. AGP9 is present along the vasculature of the reproductive tissues and AGP12 is expressed in the stigmatic cells, the chalazal and funiculus cells of the ovules, and the septum. AGP15 is expressed in all pistil tissues, except in the transmitting tract, while AGP23 is pollen grain and pollen tube specific. The expression pattern of these AGPs brings new and significant evidences for the detection of a subset of specific AGPs involved in plant reproductive processes, being of great significance for this field of study. AGPs are prominent candidates for male-female communication during reproduction.

Key words: Arabinogalactan proteins, female gametophyte, funiculus, pistil, pollen tube guidance, transmitting tract.
Introduction

All flowering plants share a common characteristic that distinguishes them from all other organisms that reproduce sexually: double fertilization (Raghavan, 2003). During this process two male sperm cells are delivered to the female gametophyte - the embryo sac -, where one fuses with the egg and the other fuses with the central cell, giving rise to the embryo and the endosperm, respectively (Russell, 1992). In order for the sperm cells to be delivered into the embryo sac several events need to occur, which implicates tightly regulated interactions between the female sporophytic tissues and the male gametophyte. Once the pollen grain is in contact with the stigmatic cells it germinates, producing the pollen tube (PT) (Kandasamy et al., 1994), which will deliver the two sperm cells to their final destination (Faure et al., 2002; Dresselhaus and Franklin-Tong, 2013). In the majority of seed plants, the PT grows through the stigmatic cells, into the style and across the extracellular matrix of the transmitting tissue in a very precise way, never losing its focus: to reach the embryo sac. Once at the funiculus, it makes a quick turn, and grows on the surface of the placenta and the funiculus until reaching the ovule opening, the micropyle (Hülskamp et al., 1995). After growing through the micropyle, the PT enters the female gametophyte, interacts with one of the two synergid cells and bursts, releasing the two sperm cells that will fuse with the central and the egg cell, ultimately giving rise to the seed and assuring the perpetuation of the next generation (Johnson and Preuss, 2002; Lord and Russell, 2002; Raghavan, 2003; Berger et al., 2008; Sprunck, 2010; Palanivelu and Tsukamoto, 2012).

During the course of all these processes, numerous cell-cell communication events must take place between different cell types. Mainly, recognition signals and attracting signals have to be sent and perceived by the female tissues and the male tissues of the plant and vice-versa, in order for a successful fertilization to occur (Dresselhaus, 2006). Until today, besides all the efforts carried out in this field of study, little information is available about which molecules function as signaling or as receptor molecules.

Arabinogalactan proteins (AGPs) constitute a large family of hydroxyproline rich proteins that are highly glycosylated and structurally complex (Showalter, 2001). AGPs are widely distributed in the plant kingdom, being ubiquitously present in land plants, also in the bryophyte Physcomitrella patens (Lee et al., 2005; Fu et al., 2007), and in all
hepatics (Basile et al., 1989), including basal angiosperms (Costa et al., 2013b) and many algae, indicating an ancient origin for these proteins (Popper et al., 2011). They are found in distinct developmental stages, cell, tissue and organ types, being mostly abundant in cell walls, plasma membranes and extracellular secretions (Majewska-Sawka and Nothnagel, 2000). AGPs are typically divided in four sub-groups according to their polypeptide core characteristics: the classical AGPs, that possess an N-terminal signal peptide, which is removed in the mature protein, a Pro/Hyp (Proline/Hydroxiproline) rich domain and a C-terminal signal for the addition of a glycosylphosphatidylinositol (GPI) anchor, the Arabinogalactan (AG) peptides, structurally similar to the classical AGPs but with a smaller protein backbone, consisting of 10 to 13 amino acid residues, the lysine-rich AGPs, with one or more lysine domains and the fasciclin-like AGPs, FLAs, with one or more fasciclin-like domains in their polypeptide core (Schultz et al., 2002; Johnson et al., 2003).

AGPs have been implicated in many important processes for plant development and growth, such as cell expansion, proliferation and differentiation, cell-cell recognition, somatic embryogenesis, PT growth, programmed cell death, seed germination and resistance to infection (Majewska-Sawka and Nothnagel, 2000). Most AGPs are predicted to be anchored to the membrane by a glycosylphosphatidylinositol (GPI) anchor (Borner et al., 2002; Schultz et al., 2004), which provides a way for the AGPs to function as signaling molecules. After comparisons with GPI anchored proteins from animal cells, two mechanisms were proposed for AGP mediated signaling: the first consisted on the cleavage of the GPI anchor by specific phospholipases (C and D) that would release the glycoprotein to the extracellular matrix, making it able to act as a signal itself or to be subject to further processing, generating different signals; the other mechanism proposed that AGPs could interact with other proteins and activate downstream signal transduction pathways (Gaspar et al., 2001; Schultz et al., 2004). Besides the hint given by the presence of the GPI anchor, implying a signaling role for these proteins, the prominent carbohydrate content surrounding the core protein also led to some assumptions about their involvement in signaling mechanisms. The importance of sugars as signaling molecules in plants is well known, and, according to some authors, the varied carbohydrate moieties of AGPs might be released via cleavage by specific enzymes (Showalter, 2001). The generated oligosaccharides might function as
signaling molecules by binding to specific membrane receptors and activating specific
signal transduction systems (Showalter, 2001). The fact that AGPs can act as chitinase
substrates, being able to stimulate somatic embryogenesis, reinforces this hypothesis,
although it is not yet demonstrated if this is an effect of the released oligosaccharides or
the modified AGP (Van Hengel et al., 2001).
AGPs have long been suggested to play important roles in sexual plant reproduction.
Earlier studies have shown the developmentally regulated enrichment of AGPs in the
extracellular matrix of the transmitting tract of several species such as Gladiolus
gandavensis, Lilium longiflorum, Nicotiana alata and Lycopersicon peruvianum
(Hoggart and Clarke, 1984; Sedgley et al., 1985; Webb and Williams, 1988, Gane et al.,
1995). AGPs have also been implicated in PT growth from the stigma to the ovules in
Amaranthus hypochondriacus, Actinidia deliciosa, Catharanthus roseus and Nicotiana
tabacum (Coimbra and Salema, 1997; Coimbra and Duarte, 2003; Cheung et al., 1995).
These studies were carried out using the β-glycosyl Yariv reagent that binds specifically
to AGPs, precipitating them (Yariv et al., 1967), or using monoclonal antibodies which
identify only the glycosidic epitopes of AGPs (Pennell et al., 1989; Pennell et al., 1991;
Knox et al., 1991). These two approaches have given us information about AGPs
distribution and localization (Coimbra et al., 2007), and clues about their possible roles
(Gao and Showalter, 2002, Sardar et al., 2006), although they allow only the detection of
general AGPs and not a specific AGP. The recent discovery that the Yariv reagent binds
specifically to the β-1,3-galactooligosaccharides of AGPs (Kitazawa et al., 2013) may
bring new insights to the possible mode of action of AGPs oligosaccharides as signaling
molecules. It will be interesting to check if this particular oligosaccharide is important
for many of the physiological processes impaired when Yariv was used in different
studies, or if Yariv only hampers AGPs ability to function by precipitating them.
Here we report the use of several constructs to explore the tissue and cell-specific
promoter activity of specific AGPs. We have focused on those AGPs, which are
particularly present along the PT pathway and other female reproductive tissues,
according to the available microarray data. With this, we aim to complement work that
has already been done by our group, describing the AGPs as molecular markers of
different stages of the Arabidopsis sexual reproductive processes (Coimbra et al., 2007).
Materials and Methods

Plant Material and Growth Conditions

Arabidopsis thaliana (L.) Heynh. seeds, ecotype Columbia were obtained from the Nottingham Arabidopsis Stock Centre (NASC). Plants were sown on soil, kept for two days at 4°C in the dark to induce stratification, and afterwards, they were grown at 22°C under a short day photoperiod (9h/15h light/dark cycles) for four weeks, followed by a long day photoperiod (16h/8h light/dark cycles) to induce flowering, with 60% relative humidity. For PAT (Phosphinotricin-Acetyltransferase) selection the seedlings were sprayed with 200 mg l\(^{-1}\) glufosinate ammonium (BASTA\(^\circledR\); Bayer Crop Science) supplemented with 0.1% Tween-20 for three or four times, every two days, during a ten day period.

Constructs generation and plant transformation

Genomic regions corresponding to the promoters of five AGPs: AGP1, AGP9, AGP12, AGP15, AGP23 were amplified using Phusion DNA polymerase (Thermo Scientific), with the primer pairs described in table 1 (supplemental material). The promoter regions were always amplified from the end of the UTR of the most proximal gene upstream of the respective AGP until its own start codon. For the genes with promoter regions with more than 3000 bp, genomic fragments of about 3000 – 3300 bp positioned upstream of the start codon of the AGP of interest were amplified. The PCR products were cloned into pENTR™/D-TOPO (Invitrogen). The resulting promoter fragments were subsequently transferred into a Gateway-compatible version (Zheng et al., 2011) of the pGreenII-based vector NLS:3GFP:NOS\(_t\) (Takada and Jürgens, 2007), termed pGII_GW:NLS:3GFP:NOS\(_t\). For AGP1, AGP15 and AGP23 GUS constructs, the respective promoter fragments were cloned into the binary vector pBGWFS7 (Karimi et al., 2002). All constructs were confirmed by DNA sequencing. The pGreenII-based expression vectors were introduced into Agrobacterium tumefaciens GV3101 harboring the pGreenII helper plasmid, pSOUP. All the others expression vectors were delivered into Agrobacterium tumefaciens GV3101 (pMP90RK). All of them were then used to transform Arabidopsis thaliana (Col-0) by the floral dip method (Clough and Bent, 1998).
Preparation of plant material for microscopy

Pistils kept in 50 mM sodium phosphate buffer (pH 7.5) were dissected under a stereomicroscope (Nikon, Model C-DSD230) by using hypodermic needles (0.4 x 20 mm, Braun). The opened carpels and the ovules that remained attached to the septum were maintained in mounting medium and covered with a cover slip.

Confocal Laser Scanning Microscopy (CLSM)

A Zeiss Axiovert 200M inverted microscope equipped with a confocal laser scanning module (LSM 510 META) was used for CLSM. GFP was excited by 488 nm and detected with a BP 505-550 filter. Optical sections were generally between 0.40 and 0.50 μm each, observed at 20x, 40x or 63x magnifications. Histology mounting medium Fluoroshield™ with 4',6-diamidino-2-phenylindole, DAPI (F6057 SIGMA) was used in order to detect the nuclei in the pollen grains. Images were captured and processed using the AxioCam HRc camera, the Zeiss LSM 510 META software and the Zeiss LSM image browser version 3.5.0.359.

Detection of GUS activity

GUS assays were performed on inflorescences as described in Liljegren et al. (2000), overnight. After chemical GUS detection, the samples were incubated in clearing solution (160 g of chloral hydrate (C-8383; Sigma-Aldrich), 100 mL of water, and 50 mL of glycerol) and incubated at 4°C overnight. The day after, inflorescences were dissected under a stereomicroscope (Nikon, Model C-DSD230), to be observed at the microscope. A Zeiss AxioImager AZ microscope equipped with Differential Interference Contrast (DIC) optics was used. Images were captured with a Zeiss Axiocam MRc3 camera using the Zen Imaging Software.

Phylogenetic Analysis

To generate the phylogenetic tree for the AGP genes, the amino acid sequences of AGPs coding sequences were aligned using Clustal W (Thompson et al., 1994) and manually edited using Jalview to reduce gaps (Clamp et al., 2004). A neighbor-joining (NJ)
(Saitou et al., 1987) tree was generated using the MEGA4 program (Tamura et al., 2007). The bootstrap values were obtained by 10,000 repetitions.

Preparation of plant material for RNA extraction

*Arabidopsis* pistils from wild-type plants were emasculated one day before anthesis and collected two days after the emasculation procedure. Pollen from *Arabidopsis* wild-type recently opened flowers was collected according to Costa et al. (2013a). *Arabidopsis* seeds were sown in half strength Murashige and Skoog (MS) medium, complemented with 0.7% agar. Agar plates were kept for two days at 4°C in the dark, to induce stratification, and subsequently they were transferred to a growth chamber at 22°C under a long day regime (16 h light/8 h dark), with irradiance of 130 μmol m-2 s-1 and 60% relative humidity. Seedlings were collected four to five days after germination.

RNA extraction, cDNA synthesis and Real Time RT-PCR

Total RNA from emasculated pistils, pollen and seedlings was extracted using PureZol™ RNA Isolation Reagent (Bio-Rad, USA) following the manufacturer’s instructions. DNA was removed by a DNase (Thermo Scientific) treatment. The isolated RNA samples were reverse transcribed using RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific) and oligo(dT)18 primers to initiate the reactions, following the manufacturer’s instructions.

cDNA was amplified using the SSoFast™ SYBR® Green Supermix on the iQ5™ Real-Time PCR Detection System (Bio-Rad) using the primers listed in table 2 (supplemental material). Real-time RT-PCR reactions were run in duplicates. After 3 min at 95 °C followed by a 10 s denaturation step at 95 °C, samples were run for 40 cycles of 10 s at 95 °C and 30 s at 60 °C. After each run, a dissociation curve was acquired to check for amplification specificity by heating the samples from 60 °C to 95 °C. Serial dilutions of pure genomic DNA from *Arabidopsis* ecotype Columbia were used to set up a calibration curve, which was used to quantify plant DNA in each sample. At the end of the PCR cycles, data were analyzed with the iQ5 2.0, Standard Edition Optical System Software v2.0.148.060623 (Bio-Rad).

Fluorescence *in situ* hybridization (FISH)
gDNA was obtained as described by Edwards et al. (1991) and used to amplify the in situ sense and anti-sense probes for AGP1 and AGP12 using the following primers:

AGP1-F 5´-CAAAAACACTCCAAAACAAA-3´, AGP1-R 5´-CTTCAGTCGGAGAATCGG-3´, AGP12-F 5´-CACAACTCATTCGACCAAAG-3´ and AGP12-R 5´-GCATCGGAAGTAGGACTTGG-3´. The amplified fragments were cloned in pGEMT-Easy (Promega). DIG-RNA probes were generated by in vitro transcription using the DIG-RNA labelling kit (Roche). The dissected pistils were permeabilized by first dehydrating in a methanol series of increasing concentration and then rehydrating in a methanol series of decreasing concentration. The pistils where then treated with 2% cellulase (Onozuka R-10) for 1 h, afterwards washed and dried. RNA/RNA fluorescence in situ hybridization was performed as described in Testillano and Risueño (2009); using DIG-RNA probes diluted 1/50 in hybridization buffer at 50 °C overnight. Post-hybridization washes were performed in 4× SSC (Saline-Sodium Citrate buffer), 2× SSC, and 0.1× SSC. Hybridization signal was detected by incubation with mouse anti-digoxigenin antibodies (1:5000 in 1% BSA, Sigma) for 90 min, followed by an incubation with Alexa-Fluor-488 anti-mouse antibody (1:25 in PBS for 45 min, Molecular Probes). After washing in PBS, sections were counterstained with DAPI, mounted in Mowiol, and observed by confocal microscopy. Controls were performed using the sense probes.

Results

Phylogenetic analysis and AGPs distribution across the genome

An alignment of full-length predicted AGP proteins was generated using Clustal W (Thompson et al., 1994) and afterwards manually refined (Fig. 1A). In this study 13 classical AGPs (AGP1, AGP2, AGP3, AGP4, AGP5, AGP6, AGP7, AGP9, AGP10, AGP11, AGP25, AGP26, AGP27), ten AG peptides (AGP12, AGP13, AGP14, AGP15, AGP16, AGP20, AGP21, AGP22, AGP23, AGP24) and three lysine-rich AGPs (AGP17, AGP18, AGP19) were considered. For this analysis only four Fasciclin-Like AGPs (FLAs) were used – FLA18, FLA16, FLA17 and FLA15. These FLAs were chosen randomly and included in the analyses only as outgroup, since they are
particularly different from the rest of the family and considered to be chimeric AGPs (Showalter et al., 2010). The phylogenetic distribution of the selected AGP sequences partially supports the four sub-groups of AGPs proposed by previous studies (Schultz et al., 2002; Johnson et al., 2003). The alignments showed a high level of similarity between the predicted amino acid sequences of AGP15, an AG peptide, and AGP1, a classical AGP, as well as the inclusion of the three lysine-rich AGPs in the same branch as the classical AGPs, not supporting the AGPs classification currently in use. As it was expected, the FLAs used in this study aligned together and independently from the other AGPs as a sub-group, but still related to the classical AGP25, AGP26 and AGP27.

Looking at the AGP gene distribution along the different five Arabidopsis chromosomes (Fig.1 B), there is no evidence of clustering of any specific group of closely related AGP genes, or any specific class of AGPs. They seem to be randomly distributed across the different Arabidopsis chromosomes.

AGPs gene expression

As a first approach, data from microarray experiments available from on-line databases such as Genevestigator (http://genevestigator.ethz.ch; Zimmermann et al., 2004) and the Arabidopsis eFPBrowser (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi; Winter et al., 2007) were used to evaluate the distinct levels of AGPs genes expression throughout the different plant tissues (Fig. 1C, only the Genevestigator data is shown). Eleven AGPs were selected for further analysis: AGP1, AGP4, AGP7, AGP9, AGP10, AGP12, AGP15, AGP16, AGP23, AGP25 and AGP26, most of them based on the presence of their transcripts in pistil tissues and their absence in stamen tissues. In the case of AGP7, although it does not show this pattern of expression, it was selected anyway, based on its predicted amino acid sequence high level of similarity with AGP4. AGP23 was selected as a negative control, since eFP Browser and the literature data (Costa et al., 2013c; Nguema-Ona et al., 2013) indicates it is expressed only in pollen. However Genevestigator data indicates a poor expression also in the female tissues. To check the differences between AGP genes expression levels among this tissues and to validate the microarray-based information, a Real Time RT-PCR was performed using emasculated pistils, pollen from flowers at anthesis (stage 13 according to Smyth et al., 1990) and seedlings cDNA. The results confirmed the microarray data initially considered (Fig. 2).
These analyses confirm the good quality of the microarray data. In this work the AGPs transcript levels were normalized to ACT8 and RUB1 reference gene levels, and are presented relative to the pollen transcript levels, since the main goal is to determine the AGPs genes that are more expressed in the female tissues than in the pollen. AGP7 and AGP23 are down-regulated in the pistil tissues when compared to their expression in pollen, while all the others AGPs are up-regulated. AGP10, AGP12 and AGP16 are the ones that revealed a higher level of over-expression when compared to their expression in pollen. AGP1, AGP4, AGP15, AGP25 and AGP26 revealed to be up-regulated in the pistils, comparing to pollen, but not in such high levels like the previous ones (AGP10, AGP12 and AGP16). From this group of up-regulated AGPs, AGP1, AGP9, AGP12 and AGP15 were selected for further analyses.

Plasmid construction and expression in A. thaliana

To improve the visualization and to avoid diffuse fluorescent signals in the detection of the promoter activities, the reporter gene NLS:3GFP was used (Takada and Jurgens, 2007). This consists of the SV40 nuclear localization signal (NLS) and three tandem enhanced green fluorescence protein (3xEGFP) sequences (Fig. 3A). The fluorescent signal should be then targeted to the nuclei, thereby enhancing the sensitivity of the GFP signal. In all the transgenic A. thaliana plants bearing the different pAGP:NLS:3GFP constructs, the GFP reporter expression has been limited to the nuclei like it was expected, as shown in Figs. 3B - M.

AGPs differential expression pattern in A. thaliana reproductive tissues

The AGP promoters selected for this study allowed us to detect the different patterns of expression of these proteins in the female reproductive tissues. All the flowers analyzed were in between stages 12 and 13 according to Smyth et al. (1990). GFP expression driven by the AGP1 promoter was strong in the style tissues (Fig. 3B), the septum (Fig. 3C), the transmitting tract (Fig. 3D), the funiculus that attaches the ovules to the placenta (Figs. 3C, D) and in the chalazal region of the ovules (Fig. 3D). A weaker GFP expression was detected in the stigmatic cells (Fig. 3B) and in the integuments of the ovule (Fig. 3D). The AGP12 promoter guided the expression of GFP strongly to the stigmatic cells (Fig. 3E) and to the chalazal pole of the ovules (Fig. 3F). A very weak
GFP expression was observed along the internal tissues of the funiculus and the septum (Fig. 3F). Plants transgenic for the pAGP15:NLS:3GFP expression cassette exhibited GFP expression in all the female reproductive tissues, except in the transmitting tract cells (Figs. 3G, H). The AGP23 promoter drove the GFP expression specifically into the vegetative cell of the pollen grains (Fig. 3I, J). This was clarified by the DAPI staining of the pollen grains, showing that the GFP signal was present only in the nucleus of the vegetative cell and not in the generative cell nuclei, where solely DAPI stained and without any green signal. The AGP9 promoter led to the expression of GFP in the vascular tissues of the: pistil transmitting tract, septum (Fig. 3K) and the funiculus (Fig. 3L), exhibiting a very weak expression in the chalazal pole of the ovules (Fig. 3M).

At the same time pAGP:GUS constructs were also analyzed for three AGPs: AGP1, AGP15 and AGP23. For the pAGP1:GUS fusion expressing plants, a low GUS activity was observed in the stigmatic cells, while a higher GUS activity was detected in the septum, transmitting tract, funiculus, chalaza and ovule integument cells (Figs. 4A, B). Regarding the plants expressing GUS under the control of the AGP15 promoter, a high GUS activity was detected almost through all the tissues of the pistil, except in the transmitting tract (Figs. 4C, D). As well as the plants expressing the three GFP molecules under the control of the AGP23 promoter, the Arabidopsis plants bearing the GUS under the control of this same promoter showed a very specific and high GUS activity in the pollen (Figs. 4E, F). This activity was also observed in the PTs (Fig. 4G), and it was especially high when the PT burst occurred inside the embryo sac (Fig. 4H), staining almost all the embryo sac with a weaker GUS signal. This GUS expression in the embryo sac was never observed when pAGP23:GUS pistils were pollinated with wild-type pollen, but only in embryo sacs fertilized with pAGP23:GUS pollen. This indicates that the GUS product present in the maternal embryo sac after fertilization is released by the burst of the PT.

FISH confirms the GFP reporter lines patterns of expression

Fluorescence in situ hybridization was used to verify if the GFP signals and GUS activity obtained with the pAGP:3GFP or pAGP:GUS fusions reflected in fact the real AGP gene expression. For this study FISH was analyzed for two AGP genes: AGP1 and AGP12. Hybridization signals for the AGP1 anti-sense probe were detected throughout
the septum, the transmitting tract and the funiculus cells as well as in the integuments surrounding the micropylar region of the embryo sac (Fig. 5A). The same experiment using the AGP1 sense probe revealed the absence of hybridization signal along all the reproductive tissues (Fig. 5B). With the AGP12 anti-sense probe strong hybridization signals were detected in the stigmatic cells (Fig. 5C) and a weaker signal was observed across the style and the septum (Fig. 5E). The corresponding AGP12 sense probe did not show any hybridization signals along the reproductive tissues (Figs. 5D, F).

Discussion

AGPs selection

Bioinformatics analyses recently allowed the identification of 64 potential AGPs in Arabidopsis (Showalter et al., 2010). The present work started by analyzing 26 of them, the ones with more information available. Even though, it is important to keep in mind that for all individual AGPs almost no information is available at the structural level. Sequence comparisons revealed a high level of similarity between amino acid sequences of AGP15, an AG peptide, and AGP1, a classical AGP, as well as the inclusion of the three lysine-rich AGPs in the same branch as the classical AGPs. These results pinpoint the artificial nature of the classification currently in use to organize this family of proteins. The availability of more data regarding AGPs expression patterns in different plant species, and more information regarding their functions may turn out possible to classify these proteins based on their functions and localization, rather than on their amino acid sequences similarities. However, there are still some pairs of AGPs that share a high degree of similarity between their aminoacidic sequences, and, simultaneously, display a similar expression pattern in the reproductive tissues, suggesting that they might act redundantly, such as for the case of AGP16/AGP20, AGP1/AGP15, AGP5/AGP10 and AGP6/AGP11 pairs. AGP6/AGP11 is a pair of redundant AGPs involved in Arabidopsis pollen grain and PT growth and development (Coimbra et al., 2009). A total of 11 AGPs were picked for further analysis: AGP1, AGP4, AGP7, AGP9, AGP10, AGP12, AGP15, AGP16, AGP23, AGP25 and AGP26. This group was selected by in silico search of AGP genes that could be transcribed preferentially in pistils rather than in the stamens or seedlings. This selection was based
on analyses of microarray data available for pistil and stamen tissues obtained from Genevestigator, using the Anatomy tool provided by this service (Zimmermann et al., 2004) and eFP Browser (Winter et al., 2007). Although AGP18 fits perfectly into this category, it was not selected as it is already well described (Acosta-García and Vielle-Calzada, 2004; Demesa-Arévalo and Vielle-Calzada, 2013). AGP23 was chosen as a control, since it is only transcribed during pollen development (Costa et al., 2013c; Nguema-Ona et al., 2013). Although microarray data from Genevestigator also predicts its expression in whole flowers and pistils, our qRT-PCR data confirmed that AGP23 is detected only in pollen, being highly down-regulated in pistils and seedlings.

The validation of this selection through qRT-PCR allowed us to limit the number of AGPs selected for further analysis to four: AGP1, AGP9, AGP12 and AGP15. AGP9 and AGP15 are up-regulated in the pistil and down-regulated in the seedlings, being selected for this reason. AGP1 was also selected, even if its transcripts show a higher up-regulation in the seedlings than in the emasculated pistils, because it’s phylogenetically close to AGP15. AGP12 was chosen as one of the most up-regulated AGPs in the pistil.

Regarding the AGP gene localization in the Arabidopsis chromosomes, it is clear that AGPs are randomly distributed over the Arabidopsis genome. This is the case of AGP16 and AGP20, contained respectively on chromosome 2 and 3. Also for AGP6, located on chromosome 5 and AGP11, on chromosome 3, two AGPs that were already shown to act redundantly (Coimbra et al., 2009). This is probably due to duplications in the genome, since most of these genes are included in segments of the respective chromosomes that were subject to large duplications events (Blanc et al., 2000). This is consistent with the prediction that genetic redundancy may occur as a consequence of gene duplication (Kafri et al., 2009). Only the pairs of most similar AGPs AGP4/AGP7 and AGP11/AGP15 have their genes positioned in the same chromosome, but in opposed regions. It is plausible that some of the AGP genes acquired a certain degree of specialization, being now expressed in different tissues, under different conditions.

AGPs expression in the reproductive tissues

The results obtained in this work confirm the specific and differential pattern of expression of AGPs previously predicted by immunolocalization studies, where several monoclonal antibodies, which recognize distinctive AGPs glycosidic epitopes, revealed
the presence of these proteins throughout diverse tissues in different developmental stages in *Arabidopsis* (Coimbra *et al.*, 2007). These results not only confirm and complement this older study but also improve the information already available about the AGPs distribution through the reproductive tissues by identifying specific AGPs present along these tissues. In Coimbra *et al.*, 2007, no antibody labelling was detected in the stigmatic cells, which, as it is shown here, are rich at least in AGP1 and AGP12. Neither it was detected any labelling in the funiculus of the ovules, whereas, in this study it is revealed the presence of several AGPs, such as AGP1, AGP12 and AGP15. This work reinforces the power of these techniques over the use of monoclonal antibodies to detect AGPs. As expected from qPCR data, *AGP23* is expressed only in pollen grains and PTs. Although microarray data available from Genevestigator expected *AGP23* to be present in pistils, this is not observed here. The analysis of transgenic Arabidopsis plants carrying the pAG23:GUS and the pAGP23:NLS:3GFP constructs revealed that both reporters were detected in pollen, proving that AGP23 is specific to the pollen vegetative cell. The prediction of potential of expression of *AGP23* in flowers and pistils is most probably due to the high levels of *AGP23* expression in pollen grains contained in the samples used for those studies. Concerning the pistil, the manipulation of these tissues is complicated if the flowers are not in the correct stage of development, being easy to get pollen contamination in the stigma, misleading into some false positive expression. A summary of all the different approaches used to localize these AGPs and their differential pattern of expression along the reproductive tissues is shown in Fig. 6.

The FISH data obtained for *AGP1* and *AGP12* are partially consistent with the promoter analysis results shown for these two AGPs. The GFP expression driven by *AGP1* and *AGP12* promoters revealed the presence of GFP signal in the chalazal tissues of the ovule, and, surprisingly, this was not observed in FISH results. This technique implies the analysis of whole ovule amounts, making the tissue permeabilization more difficult in order for the probe to reach the most internal cell layers of the ovules, as is the case of the chalazal region (García-Aguilar *et al.*, 2005; Hejátko *et al.*, 2006). Still, we are aware that maybe some regulatory elements of these two promoters might be missing, thus leading to the AGP misexpression in the chalazal tissues. Besides having regulatory sequences within the promoter itself, in eukaryotes, there may be regulatory elements
located tens thousands of base pairs away from the start site, in introns or even downstream the coding sequence of the gene (Korkuć et al., 2014). Also AGP1 transcripts were not detected in the stigmatic cells or in the style by FISH analysis. It is important to underline the fact that the microarray data used and the FISH technique were performed with whole organs, while the promoter analysis refers to a spatial-temporal analysis, much more detailed. The older immunolocalization studies (Junqueira, 2007) never detected the glycosidic AGP epitopes in these chalazal tissues. Even though we are aware that the antibodies used identify only sugar epitopes from all AGPs, we may conclude, with some cautions, that the accordance between the immunolocalization data and FISH results, fortifies the confidence in the use of antibodies to determine AGPs localization.

AGP1 and AGP12 expression in the stigmatic cells suggests the possible involvement of this protein in pollen–stigmatic cells interactions, and acquisition of pollen grain competence to initiate PT growth. Losada and Herrero (2012) pointed out a role for AGPs in supporting PT germination, suggesting that the secretion of AGPs can be associated with the acquisition of stigma receptivity in apple flower. The same mechanism may occur with AGP12 and AGP1 in Arabidopsis. Also, in the early divergent angiosperm Trithuria, immunocytochemistry results suggest AGPs to be involved in attracting the PTs through the stigmatic cuticle, as in most evolved angiosperms (Prychid et al., 2011), reinforcing our hypothesis.

The presence of AGP1 and AGP15 in the main female reproductive tissues through which the PT grows until it reaches the embryo sac – stigma, style, transmitting tract, septum, and funiculus – strengthens the putative role of AGPs in PT growth and fertilization. Many early studies implied AGPs from the female tissues as playing major roles in reproductive processes (Du et al., 1994; Cheung et al., 1995; Cheung and Wu, 1999; Wu et al., 2000; Coimbra et al., 2007). For example, TTS proteins, AGPs from Nicotiana tabacum, were shown to attract and promote PT growth either in vivo or in vitro, nutritionally supporting its growth and providing it with guidance cues (Cheung et al., 1995; Wu et al., 2000). Wu et al. (1995) also revealed that the carbohydrate part of these TTS proteins form an increasing gradient from the top to the bottom of the Nicotiana style, by the action of specific PT hydrolases, which may have a chemotropic effect on growing PTs. In Arabidopsis, the transmitting tract begins at the style between
the stigma-style boundaries, extending until the base of the ovary (Crawford and Yanofsky, 2008). AGP1 is mainly present along this transmitting tract, while AGP15 is mostly present at the septum surrounding the transmitting tract. Since these two proteins are closely related to each other, this fortifies their possible redundant function in these tissues. agp1 null mutants were analyzed (data not shown) but revealed no visible phenotype. Most probably a double agp1agp15 mutant is needed to access their precise function. These AGPs might act in these tissues in a similar manner as the TTS proteins in Nicotiana. The study of the NTT gene in Arabidopsis has indirectly implied the involvement of AGPs in PT guidance through the transmitting tract (Crawford et al., 2007). The ntt mutants lacked a functional transmitting tract and exhibited a reduced staining for acidic polysaccharides. Crawford et al., 2007 speculate that AGPs, acidic glycoproteins that are a main component of the transmitting tract, might be reduced in these mutants. It will be extremely interesting to check if there is a control of AGPs expression by this NTT zinc finger transcription factor.

AGP1, AGP9 and more strongly AGP12, show expression at the chalazal tissues of Arabidopsis ovules and at the cells located on the top of the vascular supply coming from the funiculus, as well as along this tissue. It is known that the main nutrient uptake into the endosperm occurs via the chalazal pole, being this, important for nutrient transfer from the maternal parent to the developing embryo (Debeaujon et al., 2003; Ingram, 2010). This may indicate the possible participation of these glycoproteins in nutrition or signaling between the vasculature and the embryo sac, endosperm or embryo, being quickly mobilized. The incomplete correlation between the GFP and GUS activity driven by AGP12 and AGP1 promoters in this region and their transcript expression reveals the importance of analyzing, hereafter, these AGPs at the protein level.

For double fertilization to take place, the PT must travel a long and challenging pathway, in order to reach its final destination: the micropylar entrance into the embryo sac, where it will discharge, through one of the two synergids, two immotile sperm cells to fertilize the egg cell and the central cell, giving rise to the embryo and the endosperm, respectively, initiating a whole new generation (Márton and Dresselhaus, 2010). Along this narrow road the PT lengthens through a mucilage rich extracellular matrix (ECM) from the stigmatic cells, along the specialized transmitting tract cells, the funiculus and
the ovary integuments (Webb and Williams, 1988; Lennon et al., 1998). Although most
of these studies showed that this ECM tract, through which the PT travel, is rich in
AGPs and pectins, to date, only some specific molecules have been shown to function as
PT growth enhancers such as GABA in Arabidopsis (Palanivelu et al., 2003) and
chemocyanin in Lilium longiflorum (Kim et al., 2003).
The results showed in this study, undoubtedly supports older works where AGPs were
proposed to be part of this pathway and to sustain PT growth (Clarke et al., 1979;
Herrero and Dickinson, 1979; Gell et al., 1986; Cheung et al., 1995). AGP1, AGP12 and
AGP15 (Fig. 6) are located along all these tissues and might as well contribute to PT
growth from the top of the stigma to the base of the pistil, into the ovules, either by
nutritionally supporting their growth, facilitating their movement, guiding them to their
targets or even by making them competent for PT reception by the embryo sac. These
hypotheses needs further studies to fully assign AGPs functions in these tissues, much
probably involving the obtainment of double or triple null mutants. It is interesting to
note that we identified AGPs along the entire PT pathway (stigma, style and transmitting
tract) showing that AGPs are most probably essential for all the different steps of PT
growth through the pistil. The AGPs molecular mechanism of action and how they
interact with other cell wall and cell components is still elusive, although some
enlightening has been recently given to this matter (Costa et al., 2013a). One possibility
may be related to the most recent finding that AGPs can act as calcium reservoirs,
making it available in a developmentally and temporarily way (Lamport and Vármai,
2012). It is well known the importance of calcium in sexual plant reproduction (Ge et
al., 2007). One of the key characteristics of growing PTs is a tip-focused calcium
gradient maintained by the influx of extracellular calcium through calcium channels
active at the extreme end of the growing tip (Feijó et al., 1995). AGPs may be regulating
in some way the release of calcium along the PT pathway making it available for the
PTs to grow. Most likely, different AGPs play several different roles during different
steps of the reproductive process, according to their localization and timing of
expression (Fig. 6). This work supports and improves the study of these enigmatic and
inscrutable glycoproteins in the sexual plant reproductive process, opening doors for
new pathways to the study of specific AGPs. Also, this type of analysis overcomes the
main difficulty regarding the older immunolocalization AGPs studies made by the use of
monoclonal antibodies that detect only the glycosidic epitope of the AGPs, allowing now the identification of a specific AGP in plant tissues.

Acknowledgements

This work was financed by FEDER through the COMPETE programme, and by Portuguese National funds through FCT – Fundação para a Ciência e Tecnologia (Project PTDC/AGR-GPL/115358/2009) and from an FCT PhD grant SFRH/BD/60995/2009 awarded to A.M.P. This project also benefited from financial support from the COST Action FA0903: “Harnessing Plant Reproduction for Crop Improvement”. We would like to thank Mily Ron from UC-Berkeley, Plant Gene Expression Center, for kindly sharing with us the pGII_GW:NLS:3GFP:NOST destination vector.
References


Coimbra S, Duarte C. 2003. Arabinogalactan proteins may facilitate the movement of pollen tubes from the stigma to the ovules in Actinidia deliciosa and Amaranthus hypocondriacus. Euphytica. 133, 171-178.


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

Du H, Simpson RJ, Moritz RL, Clarke AE, Bacic A. 1994. Isolation of the protein backbone of an arabinogalactan-protein from the styles of Nicotiana alata and characterization of a corresponding cDNA. The Plant Cell. 6(11), 1643-1653.


Junqueira V. 2007. *Imunolocalização de proteínas arabinogalactânicas no gineceu de Arabidopsis thaliana Wt e nos mutantes murl, mur4 e reb1-1*. Master thesis. Faculty of Sciences, University of Porto, Portugal.


The AGP protein family, gene expression and AGP gene localization in A. thaliana. (A) Phylogenetic analysis of the AGP family in A. thaliana. To generate the phylogenetic tree for AGPs, all the amino acid sequences of AGPs coding sequences were aligned using Clustal W and manually edited using Jalview to reduce gaps. Neighbor-joining (NJ) tree was generated using the MEGA4 program. The optimal tree with the sum of branch length = 14.47033254 is shown. The confidence probability (multiplied by 100) that the interior branch length is greater than 0, as estimated using the bootstrap test (10000 replicates is shown next to the branches). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. The analysis involved 30 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 241 positions in the final dataset. AGPs selected for further analysis are indicated by a yellow circle (classical AGPs) and a violet square (AG peptides). (B) The 26 AGP and the four FLA genes were localized in the Arabidopsis’ chromosomes using the Chromosome Map Tool available at The Arabidopsis Information Resource, TAIR (http://www.arabidopsis.org/jsp/ChromosomeMap/tool.jsp). (C) Gene expression pattern for the 26 AGP and the four FLA genes was obtained using Genevestigator.

qPCR relative expression levels of the selected AGPs mRNA transcripts in emasculated pistils, pollen and seedlings of wild-type Arabidopsis plants. The pollen was collected from anthers at stage 12 of flower development according to Smyth et al. (1990). AGPs transcript levels were normalized to ACT8 and RUB1 reference genes levels, and are presented relative to the pollen transcript levels. In the panel, each bar represents an average of two independent reactions and technical replicates.

Schematic representation of the expression cassette used in this study, and the resulting GFP signal shown in Arabidopsis reproductive tissues. (A) Expression cassette showing the relative position of promoter sequences (pAGP), nuclear localization signal (NLS), a fusion of three green fluorescent protein (3GFP) and the terminator Nos (NosT). (B) – (D) NLS-3GFP expression driven by the AGP1 promoter in the style
tissues (B), in the opened pistil, in the funiculus and septum tissues (C), and seen in
more detail in the transmitting tissue, funiculus and the chalazal pole of the ovule (D).
(E) – (F) NLS-3GFP expression under the control of the AGP12 promoter is observed in
the stigmatic cells (E) and in the chalazal pole of the ovule (F). (G) – (H) NL-3GFP
expression driven by the AGP15 promoter is detected in the ovule integuments, the
funiculus and the septum, but absent from the transmitting tissue (G). In (H) the GFP
signal is seen in more detail in the nuclei of the funiculus. (I) – (J) NLS-3GFP under the
control of the AGP23 promoter is absent in all the sporophytic tissues (I) being its
expression restricted to the pollen grain, and, as can be seen in the detail in (J) DAPI
staining (here in magenta) revealed this expression to be limited to the vegetative cell of
the pollen grain; DAPI stained germinative nuclei are visible (white arrowheads). (K) –
(M) NLS-3GFP signals expressed by the AGP9 promoter. Signals are observed in the
vascular bundle of the transmitting tract (K) and the funiculus (L) as well as in the
chalazal pole of the ovule (M). All the flowers used in these observations were at stage
12 and stage 13 according to Smyth et al. (1990). ch – chalaza; f – funiculus; m –
micropyle region of the ovule; ov – ovule; pg – pollen grain; s – stigma; sc – stigmatic
cell; sp – septum; st – style; v – vasculature; tt – transmitting tract. Bars: 100 μm in (B)
– (G) and (I), 50 μm in (H) and (K) – (M), 20 μm in (J).

Fig. 4 Histochemical localization of GUS activity in transgenic Arabidopsis reproductive
tissues expressing the pAGP:GUS fusion genes. (A) – (B) GUS activity driven by the
AGP1 promoter is detected in the stigmatic cells (A) and the transmitting tract, the
funiculus and the integument cells (B). (C) – (D) GUS activity driven by the AGP15
promoter observed in the ovule integuments, funiculus and septum cells. (E) – (H) a
strong GUS activity driven by the AGP23 promoter was identified inside the pollen
grains (E) and (F), the growing pollen tube (G). Upon fertilization, inside the embryo sac
a strong staining is observed at the local where the pollen tube bursts (H), followed by a
weak staining that spreads inside the whole embryo sac (H). Flowers of stage 12 and
stage 13 (Smyth et al., 1990) were used in this study. ch – chalaza; es – embryo sac; f –
funiculus; ov – ovule; pg – pollen grain; pt – pollen tube; sc – stigmatic cell; sp –
septum; sy – synergid; tt – transmitting tract. Bars = 100 μm in (A) – (E) and (G) – (H),
50 μm in (F).
Fig. 5 FISH localization of AGP1 and AGP12 transcripts in *Arabidopsis* pistil tissues. Merged images of FISH signals (green) and DAPI staining of nuclei (blue) are shown. (A) AGP1 transcript levels are detected in the funiculus, the transmitting tissue and the integuments. (C) and (E) AGP12 transcripts are localized in the stigmatic cells and along the septum tissues. (B), (D) and (F) FISH controls with the sense probe for *AGP1* in ovules (B), and *AGP12* in stigma (D), and ovules (F). All the flowers used in these observations were at stage 12 and stage 13 according to Smyth *et al.* (1990). f – funiculus; i – integuments; ov – ovule; s – stigma; sp – septum; st – style. Bars: 25 μm in (A) – (B) and 75 μm in (C) – (F).

Fig. 6 A schematic representation of the reproductive structures and tissues of *Arabidopsis thaliana* and the distribution of the 5 AGPs analyzed in this study throughout the different tissues, regarding the different techniques used. GFP presence – green; GUS presence – blue; FISH positive – yellow; experiment not performed – red; absence of signal – grey. Sc – stigmatic cell; st – style; tt – transmitting tract; sp – septum; f – funiculus; ch – chalaza; i – integuments; pg – pollen grain; v – vasculature.
Figure 2

![Bar chart showing relative normalized expression of AGP1 to AGP26 in Pistil and Seedling stages.](chart.png)
Figure 3

A

- pAGP  NLS  3xGFP  NosT

B C D E

F G H I

J K L M

st  ov  sp  sc

f  ov  sp

chi  tt

m

pg

sp  v  cb

pg  ov  cr

f  ov  cr
Figure 5

A  anti-sense AGP1

ov  f  sp  tt

B  sense AGP1

ov  f  i

C  anti-sense AGP12

s  st

D  sense AGP12

pg  s  st

E  anti-sense AGP12

f  sp  ov  tt

F  sense AGP12

tt  ov  sp  f
Arabidopsis thaliana

Figure 6