Mycorrhiza (2008) 18:403–412
DOI 10.1007/s00572-008-0194-1

ORIGINAL PAPER

# Novel plant and fungal AGP-like proteins in the *Medicago truncatula–Glomus intraradices* arbuscular mycorrhizal symbiosis

**Carolyn J. Schultz · Maria J. Harrison**

Received: 27 March 2008 / Accepted: 18 July 2008 / Published online: 16 August 2008


**Abstract** The ability of arbuscular mycorrhizal (AM) fungi to colonise the root apoplast, and in coordination with the plant develop specialised plant–fungal interfaces, is key to successful symbioses. The availability of expressed sequence tags (EST) of the model legume, *Medicago truncatula*, and AM fungus, *Glomus intraradices*, permits identification of genes required for development of symbiotic interfaces. The *M. truncatula* EST database was searched to identify cell surface arabinogalactan-proteins (AGPs) expressed in mycorrhizal roots. Candidate genes were characterised and gene expression tested using reverse transcription polymerase chain reaction and promoter: reporter gene fusions. Genes encoding one plant AGP and three AGP-like (AGL) proteins (from *G. intraradices*) were identified. AGL proteins encoded by two AGL genes from *G. intraradices* (GiAGLs) represent a new structural class of AGPs not found in non-AM fungi or plants. Two GiAGLs differ from plant AGPs by containing charged repeats. Structural modelling shows that GiAGL1 can form a polyproline II helix with separate positively and negatively charged faces, whereas GiAGL3 is charged on all three faces. The unique structural properties of the newly discovered AGLs suggests that they could assist the formation of symbiotic interfaces through self-assembly and interactions with plant cell surfaces.

**Keywords** Arabinogalactan-proteins (AGPs) · Hydroxyproline-rich glycoproteins · Polyproline II helix · Glycosylphosphatidylinositol anchor · Mycorrhiza

## Introduction

In arbuscular mycorrhizal (AM) symbioses, the fungus inhabits apoplastic spaces within the root where novel interfaces are created in which walls of both symbionts are modified (Peterson and Massicotte 2004; Balestrini and Lanfranco 2006) and over which nutrients are exchanged (Harrison 2005; Smith et al. 2006). The fungal and plant proteins that control development of these interfaces are still poorly understood. The fungal cell wall at the interface is thin and contains low levels of chitin, and the plant cell wall is also reduced (Bonfante-Fasolo et al. 1990a; Peterson and Massicotte 2004).

There are two main morphological types of AM symbioses with structurally different interfaces. *Arum*-type AM are characterised by highly branched intracellular arbuscules, subtended by intercellular hyphae, whereas in *Paris*-type AM, intracellular fungal coils grow directly from cell to cell, with little or no intercellular phase (Dickson et al. 2007). The type and amount of wall material of both symbionts varies in these different interfaces. For example, *Paris*-type AM with *Glomus intraradices* and ginseng (*Panax quinquefolius* L.) produces an interface with abundant non-esterified pectin around the coils but less in the finer arbusculate coils (Armstrong and Peterson 2002). This distribution of pectin is different from that observed in *Arum*-type AM with *G. versiforme*

**Electronic supplementary material** The online version of this article (doi:10.1007/s00572-008-0194-1) contains supplementary material, which is available to authorized users.

C. J. Schultz (✉)
School of Agriculture, Food and Wine, University of Adelaide,
Glen Osmond, SA 5064, Australia
e-mail: carolyn.schultz@adelaide.edu.au

M. J. Harrison
Boyce Thompson Institute for Plant Research,
Tower Road,
Ithaca, NY 14853, USA

and leek (*Allium porrum*) where pectin can be found in the interfacial material of hyphae deep within cortical cells (Bonfante-Fasolo et al. 1990b).

Arabinogalactan-proteins (AGPs) are an important class of cell surface proteoglycans with important roles in plant development (Gaspar et al. 2001; Seifert and Roberts 2007), and they are found at AM symbiotic interfaces (Gollotte et al. 1995; Balestrini et al. 1996; Balestrini and Lanfranco 2006). AGPs are plant proteoglycans that are mostly carbohydrate (90–98%) attached to a protein backbone (2–10%) rich in hydroxyproline (Hyp), Ala, Ser and Thr. The AGP family is well characterised in *Arabidopsis thaliana* (Schultz et al. 2002; Johnson et al. 2003), but knowledge of AGPs in other plants is limited. *Arabidopsis* does not form AM symbioses; therefore, *Medicago truncatula* may contain AM-related AGPs not found in *Arabidopsis*. To identify AGPs expressed in AM symbioses, we analysed the available genome and expression data for *M. truncatula* (VandenBosch and Stacey 2003; Hohnjec et al. 2006).

## Materials and methods

### Plant and fungal material

Plant materials used in this study were *M. truncatula* cv. A17 and hairy root cultured *Daucus carota* clone DC2 (Bécard and Piché 1992). AM fungal species *G. intraradices* (DAOM 181602; St-Arnaud et al. 1996), *G. versiforme* (INVAM IT104) and *Gigaspora gigantea* (obtained from D. Douds; Nagahashi and Douds 2007) were maintained as previously described (Liu et al. 2003, 2007).

### Database searches

The protein sequence of glycosylphosphatidylinositol (GPI)-anchored classical AGPs and AG peptides from *A. thaliana* (Schultz et al. 2002) were used in Tblastn searches (at http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/Blast/index.cgi). The expression summary was evaluated to find genes expressed in AM roots.

### Reverse transcription PCR of *M. truncatula* roots

RNA was extracted from *M. truncatula* tissues using a modified Trizol (Invitrogen) method. Tissues used were roots colonised by *G. versiforme* (Liu et al. 2003), *G. intraradices* or *Gi. gigantea* (Liu et al. 2007), and their respective mock-inoculated controls, and flowers, leaves, seed pods and seeds. Approximately 80 ng of complementary DNA (cDNA) was used in 20-µl standard polymerase chain reaction (PCR) using forward (F1) and reverse (R1) primers for TC94753, TC109005 and TC103739, respectively (Table S1). Elongation factor-α1 (EF1) was used as a positive control (Table S1). Cycling conditions were 94°C 1 min, then 30 cycles of 94°C 20 s, 55°C 30 s, 72°C 30 s. Selected reactions were re-amplified and cloned into pGEM-T easy (Promega, Madison, USA) and sequenced.

### Plasmid construct and plant transformation

A 2,067-bp fragment of the *MtAMA1* promoter (BAC mth2-175p8; GenBank# CT033767) was PCR-amplified using primers AMA1_2067p_F (−2,067 bp from the translation start site; ATG) and AMA1_Pro_R (−1 bp from the translation start site; ATG; Table S1) using ExTaq polymerase (Takara, Shiga, Japan) following manufacturers' instructions. The PCR fragment was digested with Hind*III* and Xba*I* and subcloned into a binary vector pROK_MCS_GUS (pCAMBIA2301 (GenBank# AF234316) with the *Hin*dIII-CaMV 35S promoter-GUS-*Bst*EII fragment replaced by a PCR-generated promoter-less construct *Hin*dIII-GUS-*Bst*EII fragment). Cloning was confirmed by DNA sequencing, then the plasmid transformed into *Agrobacterium rhizogenes* strain Arqua1 as previously described (Boisson-Dernier et al. 2001). *M. truncatula* roots were transformed as described by Boisson-Dernier et al. (2001), inoculated with *G. versiforme* spores 2 weeks after transfer to pots, then grown for a further 4 weeks.

### Staining of transformants for GUS activity and wheat germ agglutinin

Transformed roots were stained for GUS activity with X-Gluc (5-bromo-4-chloro-3-indolyl-β-D-glucuronide) for 5 h at 37°C in the dark and then fixed and stained with wheat germ agglutinin-AlexaFluor 488 (WGA) to visualise the fungus as previously described (Liu et al. 2003).

### Genomic DNA extraction and PCR

Genomic DNA was extracted from ≈100 mg tissue (*M. truncatula* leaf and carrot hairy roots) or $1\times10^5$ fungal spores (*G. intraradices* and *G. versiforme*) by grinding in liquid nitrogen and extracting in 500-µl buffer [0.1 M Tris·Cl pH 8.0, 0.05 M ethylenediaminetetraacetic acid, 1.25% sodium dodecyl sulphate (SDS)] for 1 h, 65°C. DNA was precipitated with 0.5 volumes of 6 M ammonium acetate (4°C, 30 min). Supernatant (500 µl) was precipitated with 300 µl isopropanol. Pellet was rinsed with 70% ethanol and resuspended overnight at 4°C in 250 µl water. Touchdown PCR was performed on genomic DNA used the following primer pairs: AMA1_cons_F1, AMA1_cons_R1;

DcAGP1_F1, DcAGP1_R1; GiAGL1_degF1, GiAGL1_R2; GiAGL2_F2, GiAGL2_R2; and GvF1, GvR1 (Table S1). Cycle conditions were 92°C for 2 min, followed by 10 cycles [94°C 30 s, 68°C 30 s (decreasing by 0.8°C per cycle), 72°C 1 min], followed by 30 cycles [94°C 30 s, 56°C 30 s (decreasing by 0.8°C per cycle), 72°C 1 min].

### RT-PCR of fungal tissue and plant hairy roots

RNA was isolated (as above) from freshly harvested spores (4-month-old plate), extraradical hyphae (6-week-old plate) and roots from either mono- (7-week-old plate) or bi- (6-week-old) hairy root cultured *D. carota* clone DC2 (Bécard and Piché 1992). RNA was also isolated from spores that had been stored in water at 4°C for 4 weeks then germinated (in water) at room temperature for 48 h. All samples were treated with DNAfree (Applied Biosystems/Ambion, Texas, USA). Approximately 400 ng of cDNA was used in 25 µl standard PCR reactions using forward (F) and reverse (R) primers for *GiAGL1* (GiAGL1_F2 and GiAGL1_R2), *GiAGL2* (GiAGL2_F2 and GiAGL2_R2), *GiAGL3* (AGL3_F1 and GiAGL3_R1) and the 18S ribosomal RNA (rRNA) genes from *G. intraradices* (Gi18SrRNA_F and Gi18SrRNA_R) and carrot (Dc18S_F and Dc18S_R) (Table S1). Cycling conditions were 94°C 2 min, then 30 cycles of 94°C 20 s, 55°C 30 s, 72°C 60 s, except for Gi18S where a 58°C annealing temperature was used. PCR products for *AGL* genes were electrophoresed and transferred to nylon membrane for probing with digoxigenin-labelled PCR products as previously described (Schultz and Coruzzi 1995). Hybridisation was performed at 65°C in an aqueous hybridisation solution followed by two high stringency washes (15 min each in 0.1× SSC and 0.1% SDS). Products for rRNA genes were visualised with ethidium bromide staining and the images inverted in Adobe Photoshop.

### Identification of Pro, Ala, Gly-rich proteins in fungal genomes

Annotated fungal genomes were downloaded from http://genome.jgi-psf.org/euk_cur1.html (20th Dec 2007) or ftp://genome-ftp.stanford.edu/pub/yeast/data_download/sequence/genomic_sequence/orf_protein/ (10 Jan 2008). Proteins with 50% Pro, Ala, Gly were identified as described for *A. thaliana* AGPs (Schultz et al. 2002).

### Modelling of polyproline II helices

Four repeats each of Ala-Pro-Ala-Asp-Gly-Lys (APADGK) and Ala-Pro-Lys-Asp-Gly (APKDG) representing GiAGL1 and GiAGL3, respectively, were modelled in DeepView (http://www.expasy.org/spdbv/) as polyproline II helices by manually setting the Phi and Psi angles to −75 and +145 (Bochicchio and Tamburro 2002).

## Results

### Three AGP-like genes expressed specifically in mycorrhizal roots

Tblastn analyses of the *M. truncatula* gene index identified 20 genes that encode proteins with non-contiguous Ala-Pro, Ser-Pro and Thr-Pro repeats, an endoplasmic reticulum (ER) secretion signal and a GPI anchor signal (data not shown). Three genes, TC94753, TC109005 and TC103739, were represented by expressed sequence tags (ESTs) arising exclusively from *M. truncatula*/*G. intraradices* mycorrhizal root libraries (Table 1). RT-PCR showed that TC94753 is expressed in roots colonised by three different AM fungi, *G. versiforme*, *G. intraradices* and *Gi. gigantea* (Fig. 1). Expression of TC109005 was not detected in any of the samples, whereas TC103739 was expressed only in the roots colonised by *G. intraradices.* None of the genes were expressed in the above-ground tissues.

### MtAMA1 promoter:GUS fusions are expressed in arbuscule-containing cells

The gene corresponding to TC94753 will be referred to as *MtAMA1*, for *a*rbuscular *m*ycorrhiza *A*GP. *M. truncatula* transgenic roots expressing *MtAMA1* promoter:GUS were colonised by *G. versiforme* and GUS expression examined. Of six independent transformants, four showed positive GUS staining exclusively in cells containing arbuscules (Fig. 2a,c,e). The presence of arbuscules was confirmed by counterstaining with WGA-AlexaFluor 488 (Fig. 2b,d,f). Not all of the arbuscule-containing cells were stained with GUS (compare Fig. 2e,f). Two other independent transformants showed GUS activity in arbuscule-containing cells and minor vascular staining. Control plants, transformed with a 35S:GUS construct, showed GUS staining throughout the root (Fig. 2g,h) especially at the root tip (Fig. 2i).

### AGP-like genes in *G. intraradices*

TC109005 and TC103739 were not present in the *M. truncatula* genome sequence databases, and attempts to amplify the genes from *M. truncatula* genomic DNA were unsuccessful. Sequencing of the products from the RT-PCR reactions (*G. intraradices* colonised roots, Fig. 1) did not match the sequence expected for the contigs TC109005 and TC103739. The PCR fragment amplified with primers for

**Table 1** Protein sequences and EST expression summary

| Predicted protein sequence[a] | | cDNA libraries | | |
|---|---|---|---|---|
| | | MtBC[b] | MTGIM[c] | #9CR[d] |
| TC94753 | *MASPMKFSMVIACVLVLVMVVAA*QNNGEDGINIKVSND<br>MPGMVMAPAPTPKS*SASLPTLTYSAAILIFLPFMLSFF* | 13 | 20 | 1 |
| TC109005 | *MAKFTKLTFVCVLVLVLLVSFSVSTPELAVR*QAPGDA<br>TPVGGDATPGATPPAGGPPAGATPPAGATPPAGGPPAGA<br>TPPAGGPSSSPSGAAATSSSASAAGPSGSSTPT<br>S*AAYKIESGLSSVAALVALVGFF*L | 4 | – | – |
| TC103739 | *MKFNKRILFLLAVLTVLSAYVAA*QAPGAPAPAPGGDAGKAP<br>APAPGGDGKAPAPGGAPAGKAPAPGGDGKAPAPGG<br>APAGKAPAPGGDAGKAAPGTPPAGGPPGVTPSGSASP<br>PAATPAAAAPKSSTGATGTTAATGS*GNSLKSEFGISLAAIAVLGAIFA* | 1 | – | – |
| MtAM1[e] | *MKFNNRIIFLLVVLAVLIACVAA*QGPVGAPPAPGTPPPAEPA<br>PGAPPPPPKGKDAKGKDTPDGDAAKGKDAAPDGA<br>KGKDAAPDGAKGKDAPKEGAKGTVTPPAPAAPGAAPG<br>AAPGTAPAPGGPPPEGAAPSPAKGGAAAPTPGAGTGTSVA<br>PAGASGSTAAKTATG*AGNSLKSEVGVSFVAVILGALFA* | – | – | – |

[a] Translations and expression summaries from http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gireport.pl?gudb=medicago. The cleaved N- and C-terminal signal sequences are italicised and were predicted by http://www.cbs.dtu.dk/services/SignalP/#submission and http://mendel.imp.ac.at/gpi/plant_server.html, respectively.
[b] Harvested 3 weeks post-inoculation with *G. intraradices* (5520)
[c] Mycorrhizal roots 3 weeks after inoculation (#ARB)
[d] Mycorrhized roots 3 weeks (#9CR)
[e] MtAM1 is an AGP encoded by a cDNA clone from mycorrhizal roots of *M. truncatula* colonised by *G. versiforme* (van Buuren et al. 1999). A genomic clone corresponding to this cDNA has not yet been deposited at the National Center for Biotechnology Information databases (tBLASTN search June 2008, of non-redundant, genomic survey sequences and high-throughput genomic sequences).

TC109005 was 398 bp and approximately 75% identical to the *M. truncatula* contig (nt 1..275 were 71% identical to the cloned PCR fragment and nt 105..398 were 75% identical). The PCR fragment amplified with primers for TC103739 was 658 bp and was less than 70% identical to the *M. truncatula* contig (nt 1..394 were 60% identical and nt 248..658 were 70% identical to the cloned PCR fragment). The sequence of the product amplified with primers specific for TC103739 showed a strong match to ESTs from libraries generated from germinating spores (GenBank # BM959500 is 93% identical to nt 1..650 of the cloned PCR fragment) and extraradical hyphae of *G. intraradices*. New primers were designed and tested on genomic DNA from *M. truncatula*, carrot, *G. intraradices* and *G. versiforme*. The primers successfully amplified the two genes from *G. intraradices* genomic DNA, but not the other DNA samples tested (Fig. 3a). Primers for DcAGP1 and "Gv+ve" were used as positive control reactions to confirm the integrity of the carrot and *G. versiforme* DNA samples. These genes are named *GiAGL1* and *GiAGL2* for *AGP*-like genes.

## Expression of *GiAGLs*

Both *GiAGL1* and *GiAGL2* are expressed in *G. intraradices*-colonised roots, but not mock-inoculated roots or

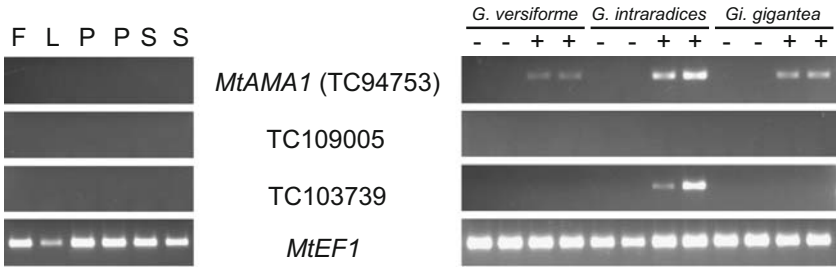



**Fig. 1** Expression of AGPs in AM colonised roots. Gene-specific primers for three *M. truncatula* EST contigs (TC94753, TC109005 and TC103739) were used in RT-PCR experiments using RNA from flowers (*F*), leaves (*L*) and two independent RNA samples of the following tissues: pods (*P*), seed (*S*) and roots samples that were colonised (+) or mock inoculated (−) with three different AM fungi, *G. versiforme*, *G. intraradices* and *Gi. gigantea*. Primers for elongation factor (EF1 were used as a control)

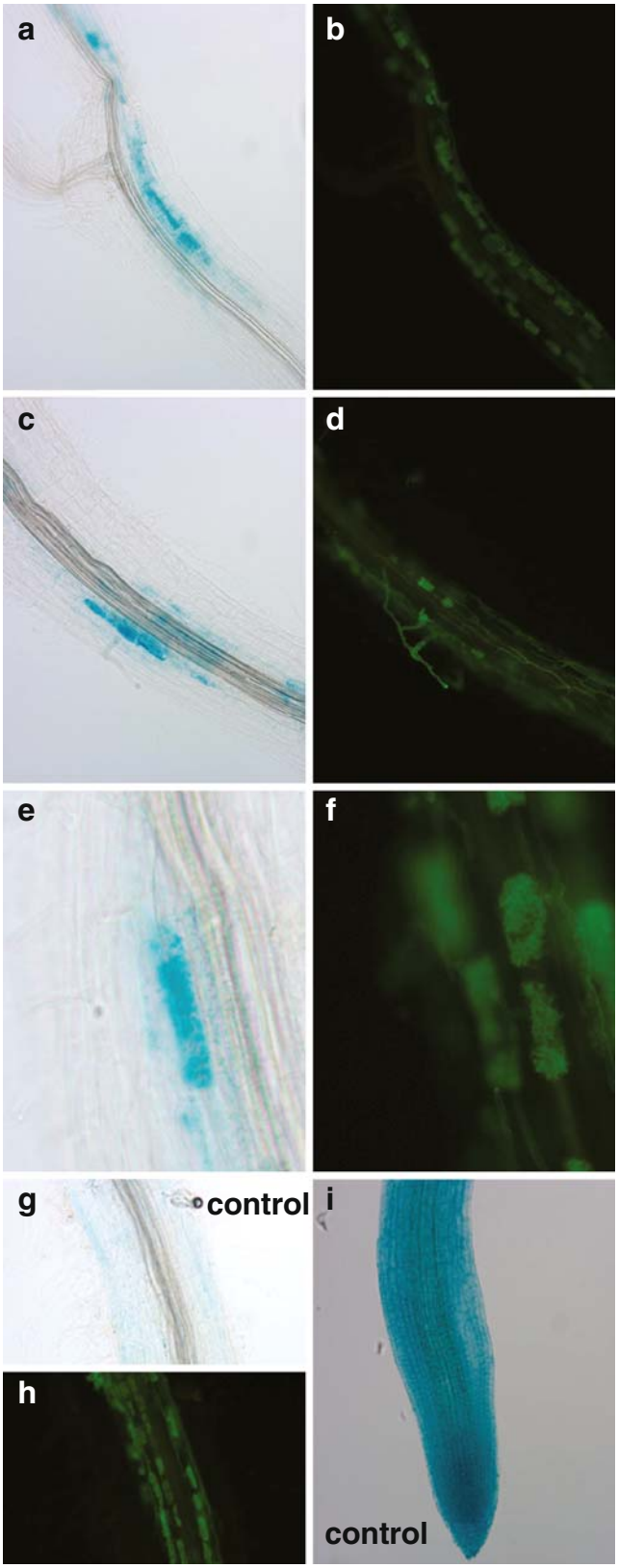



**Fig. 2** *MtAMA1* is expressed in arbuscule-containing cells. Wild-type (A17) *M. truncatula* roots transformed (*Agrobacterium rhizogenes*) with *MtAMA1* promoter:GUS constructs (**a**–**f**). Roots were stained with GUS, fixed, then co-stained with wheat germ agglutinin (*WGA*). **a**, **b** GUS and WGA staining (respectively) of the same section of colonised root from transformant 1; **c**, **d** transformant 2; **e**, **f** transformant 3. **g**–**i** Transformants with a 35S promoter:GUS construct

roots colonised by *G. versiforme* or *Gi. gigantea* (Fig. 3b). A third AGL gene, *GiAGL3*, was also discovered (as described later) and also is expressed in mycorrhizal roots colonised by *G. intraradices* but not the other two AM fungi. Since ESTS for *GiAGL1* and *GiAGL2* exist in EST libraries from germinating spores and extraradical hyphae, we examined the expression of the AGL genes in in vitro grown fungal and hairy root tissue. PCR gel blots of RT-PCR products showed expression of *GiAGL1*, *GiAGL2* and *GiAGL3* in freshly harvested spores (Fig. 3c). *GiAGL2* and *GiAGL3* are both expressed in 2-day germinating spores. None of the genes showed expression in extraradical hyphae, which may reflect the different growth conditions (phosphate treatment) of the hyphae used to generate the EST population from extraradical hyphae (Maldonado-Mendoza et al. 2004). Only *GiAGL2* was expressed in mycorrhizal carrot hairy roots, and none of the genes were expressed in non-mycorrhizal roots. 18S ribosomal rRNA genes from *G. intraradices* (Gi 18S) and carrot (Dc 18S) were used as controls. Selected no reverse transcriptase (no RT) controls were included to confirm the absence of genomic DNA (bottom three panels).

## *GiAGL1* encodes a repetitive amphipathic protein

Accurate full-length sequence for *GiAGL1* was obtained by sequencing two independent cDNA clones from a *G. intraradices* extraradical mycelium library, previously submitted as ESTs (GenBank accessions: BI452303 and BI452297). The encoded protein includes an N-terminal signal sequence and a C-terminal GPI anchor signal (Fig. 4a). GiAGL1 differs from *Arabidopsis* AGPs in three ways: (1) it contains high levels of Gly (15.0%), whereas classical plant AGPs contain between 0.5% Gly (AtAGP9) and 6.3% Gly (AtAGP5); (2) it is more repetitive than plant AGPs containing 17 repeats of Ala-Pro-Ala-Asp-Gly-Lys (APADGK) including one block of ten consecutive repeats (Fig. 4a); and (3) it contains negatively and positively charged residues, Asp (11.1%) and Lys (10.6%), whereas plant AGPs have very few charged residues within the Ala-Pro-rich region (data not shown).

The partial sequence obtained for *GiAGL2* (from the cloned RT-PCR product) predicts a protein that is expected to be GPI-anchored based on partial ER and GPI anchor signal sequences (Fig. 4b). GiAGL2 is more plant-like than GiAGL1 because it is less repetitive and does not contain charged residues, although it does have high levels of Gly (15.4%).

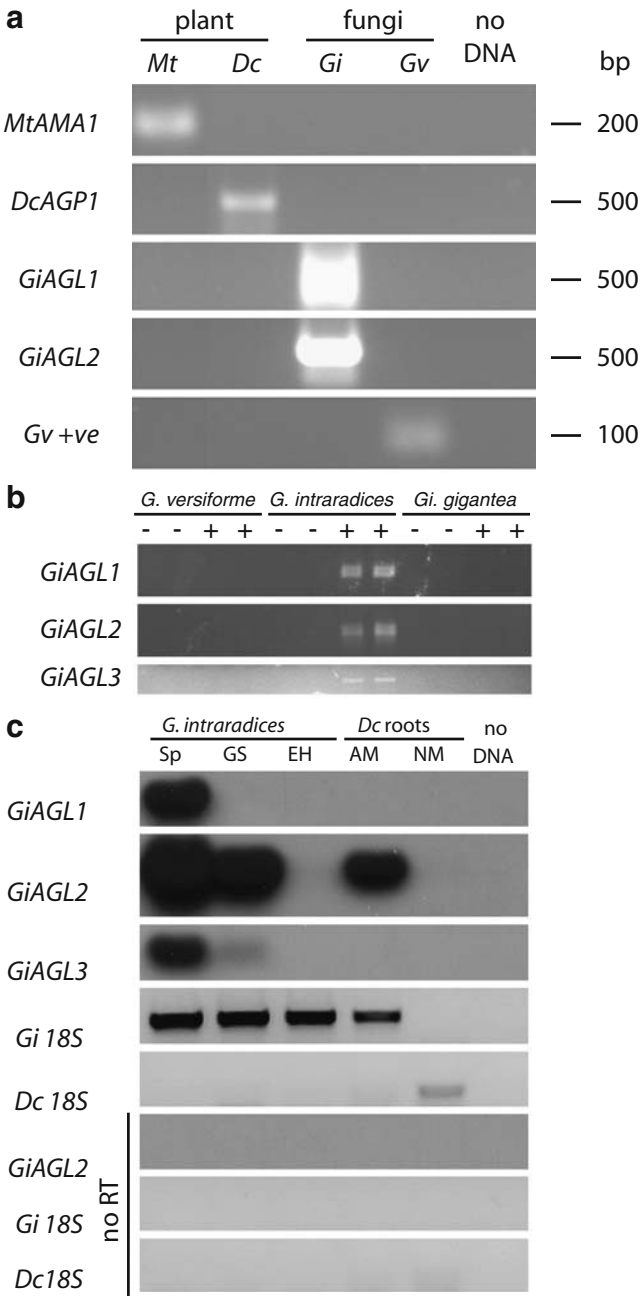



**Fig. 3** Expression of genes encoding AGP-like proteins from *G. intraradices* but not *G. versiforme*. **a** PCR from plant (*M. truncatula* and carrot) and fungal (*G. intraradices* and *G. versiforme*) genomic DNA to show the origin of AGL genes. Gene-specific primers were designed to *MtAMA1*, carrot *DcAGP1*, *GiAGL1, GiAGL2* and a *G. versiforme* positive (+ve) control gene (GenBank # CA860189). **b** RT-PCR using the same cDNA samples as used in Fig. 1, except with new primers for *GiAGL1* and *GiAGL2* based on sequence of cloned RT-PCR re-amplification products and primers designed to *GiAGL3* based on GenBank# BI452321 (Table S1). **c** RT-PCR of fungal tissue and carrot hairy roots. PCR gel blots of RT-PCR products to detect expression of *GiAGL1*, *GiAGL2* and *GiAGL3* in spores (*Sp*), germinating spores (*GS*), extraradical hyphae (*EH*) from *G. intraradices* and in mycorrhizal (*AM*) and non-mycorrhizal (*NM*) carrot hairy roots. 18S ribosomal rRNA genes from *G. intraradices* (Gi 18S) and carrot (Dc 18S) were used as controls. Selected no reverse transcriptase (*no RT*) controls were included to confirm the absence of genomic DNA (*bottom three panels*). A no DNA control was included for every primer mix to monitor PCR contamination

### GPI-anchored Pro-rich proteins are rare in fungi

An amino acid bias program (Schultz et al. 2002) was used to search for proteins with >50% Pro, Ala and Gly in the completed genomes of seven fungal species: three basidomycetes (an ectomycorrhizal fungus, *Laccaria bicolor*, the wood decomposer, *Phanerochaete chrysosporium* and the brown rot fungus, *Postia placenta*), a zygomycete (*Phycomyces blackesleeanus*) and three ascomycetes (*Aspergillus niger*, *Saccharomyces cerevisiae* and *Trichoderma reesei*). Only two of these genomes contained Pro-, Ala- and Gly-rich proteins that are predicted to be GPI-anchored (Table S2).

The protein from *L. bicolor* (Table S2) has low similarity to the two proteins from *G. intraradices*, GiAGL1 and GiAGL2. However, it has 43% similarity to a third AGP-like protein from *G. intraradices*, here designated GiAGL3 (Fig. 4c), encoded by cDNA GenBank# BI452321. GiAGL3 is more like GiAGL1 than GiAGL2, as it is repetitive and contains charged residues, although the repeats contain five amino acids Gly-Ala-Pro-Lys-Asp (GAPKD).

### Modelling of repeats in AGL1 and AGL3 reveals distinct polyproline II helices

The repetitive amphipathic (positively and negatively charged) nature of GiAGL1 and GiAGL3 suggests that they have important structural properties (Pujals et al. 2006; Rauscher et al. 2006). Structural modelling of the repeats of both proteins (Fig. 4d,e) shows that they can form polyproline II helices with distinct surface properties. A model of the AGP domain of MtAMA1 is shown for comparison (Fig. 4f).

## Discussion

### MtAMA1 is a chimeric AG peptide

MtAMA1 is a chimeric AG peptide (Johnson et al. 2003) with a unique 21 amino acid N-terminal domain followed by a putative short motif for AG glycosylation (Tan et al. 2003), Ala-Pro-Ala-Pro-Thr-Pro (Table 1). The non-contiguous Ala-Pro-Ala-Pro-Thr-Pro motif is characteristic of AGPs where Pro residues are hydroxylated then glycosylated. Orthologs of *MtAMA1* are not found in

**Fig. 4** Fungal AGL proteins contain amphipathic repetitive domains. **a**The predicted amino acid sequence of GiAGL1 (new GenBank # EU931681) obtained by sequencing two independent cDNA clones GI08_D02 (GenBank # BI452297) and GI08_G08 (GenBank # BI452303). **b** The predicted amino acid sequence of GiAGL2 obtained by RT-PCR from *G. intraradices* colonised *M. truncatula* roots (new GenBank # EU931682). The "x" residues in **b** are not known for this gene, but based on sequence similarity to the related, and apparently full-length "*M. truncatula*" EST TC109005, this predicted protein is expected to contain ER and GPI anchor signal sequences. **c** The predicted amino acid sequence of GiAGL3 based on the sequence of EST (GenBank # BI452321). The repeat regions (**a–c**) are *shaded in different coloured boxes* which are staggered to emphasise the repeats. The ER and GPI-anchor signal sequences (**a–c**) are underlined. **d** Space filling model of four APADGK repeats of GiAGL1 based on a polyproline II helix. **e** Space filling model of four APKDG repeat of GiAGL3 based on a polyproline II helix. **f** Space filling model of the AGP domain of MtAMA1. Lys residues (+) are in *pink*, Asp residues (−) are in *blue* and Pro residues in *brown*

**a** *GiAGL1*

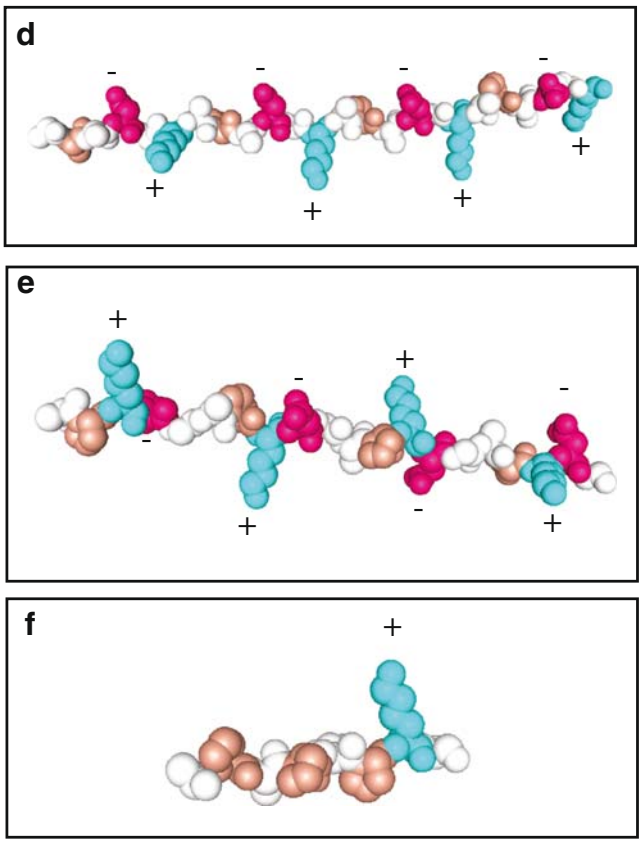
MNKRIFIILAIFFALVTYVAAQGPADG
APADGKAPADGAPADGKAPADGAPADGKAPADG
APADGKAPADGKAPADGKAPDAADGKAPADGK
APADGKAPADGKAPADGKAPADGKAPADGK
APADGKAPADGKAPADGKAPADGKAPADGK
APAGAAPGGAAPAPGAAPKDGKAPAPADGK
TPAASPAGGAAPSASAKAAA
SSGNSLKSSGYSFAAIAVLGAIFA

**b** *GiAGL2*

xxxxxxxxxxxxLVLVLLVSFSVS
SPERLAVRQAPTPDTGAPAPGGDATPPAGGATPGATPPAA
GGAPPAGATPPAAGGAPPAGATPPAGATPPAGGAPPAGGP
SSSPSSSPTVSASAAGPSGSSSPAATG
AAYKIESGLSSVAAxxxxxxxxx

**c** *GiAGL3*

MKFNNRIFFLLIVLTVLIACVAA
QGPGGAPGAPAPAPGAPPPDGAPKDGAPKDGAPKDGAPKD
GAPKDGAPKDGAPKDGAKGAPKDAPKDAPKDAGGDAKGKV
APPAGGAPPAGGAPAGPPPEGAAPSPAKTAAPTPGGGTGT
SVAPAGASGSTPAKSATG
AGNSLKAEVGVSFAAVILGAIFA




GiAGL1 is predicted to be an elastic polymer of the fungal cell wall

GiAGL1 and GiAGL3 contain repeat domains capable of forming polyproline II helices with distinct structural features (Fig. 4d,e). Polyproline II helices are extended rods with three amino acid residues per turn, leading to three faces (Bochicchio and Tamburro 2002). In GiAGL1, the positive and negative charges are found on different

Arabidopsis or rice; however, a putative pea ortholog, represented by a single EST (GenBank# AJ308187), is also expressed in mycorrhizal roots (Grunwald et al. 2004). *MtAMA1* is different from the previously reported *MtAM1* that encodes an AGP with a protein backbone of 145 amino acid residues (after cleavage of N- and C-terminal; Table 1).

*MtAMA1* is expressed in mycorrhizal roots, and the level of expression is apparently higher in the mycorrhizal roots colonised by *G. intraradices* than it is with *G. versiforme* and *Gi. gigaspora* (Fig. 1). This could be due to the higher levels of colonisation with *G. intraradices* than the other two AM species (Liu et al. 2003, 2007) and can be tested in future studies with quantitative real-time PCR using both plant and fungal primers. Expression of GUS driven by the *MtAMA1* promoter is restricted to a subset of arbuscule-containing cells (Fig. 2e,f), suggesting that this gene is only required for a specific window of arbuscule development. A variety of techniques (e.g. RNA interference and time course experiments) will be needed to determine if this gene is required for establishment, maintenance and/or turnover of arbuscules.

It is difficult to predict the function of MtAMA1, since the non-AGP domain shows no similarity to other protein domains. The presence of the GPI anchor suggests that it could act as a co-receptor at the plasma membrane, or a potential ligand /signalling molecule, if released from the plasma membrane by phospholipase activity (Schultz et al. 1998). The small size of MtAMA1 makes it a good candidate for a signalling molecule.

faces of the helix (Fig. 4d) because each repeat (APADGK) forms two complete turns, whereas in GiAGL3, the five amino acid repeats (APKDG) result in charged residues on all three faces (Fig. 4e).

Polyproline II helices are found in naturally occurring elastic proteins such as spider webs, wheat gluten and elastin (Rauscher et al. 2006). Not all of these proteins have a polyproline motif; for example, elastin, found in arteries and skin, has a six residue motif with only a single Pro residue (PGVGVA; Rauscher et al. 2006). The percentage of Pro and Gly in repeats within GiAGL1 (17% Pro, 17% Gly) and GiAGL3 (20% Pro, 20% Gly) suggest that these repeat regions can adopt either elastic or amyloid (hard waxy deposits) properties depending on the solution conditions (Rauscher et al. 2006).

A wide range of proteins that form amphipathic polyproline II helices have cell-penetrating activity (Pujals et al. 2006). The amphipathic nature of cell penetrating peptides allows them to self-assemble, a feature that is easily imagined from the models of GiAGL1 and GiAGL3 (Fig. 4d,e). The structural flexibility of GiAGL1 and GiAGL3 suggests roles in the establishment and/or maintenance of the plant–fungi interface. When *G. intraradices* colonises *M. truncatula*, it forms highly branched *Arum*-type arbuscules (Dickson et al. 2007), resulting in extension of the plant plasma membrane. We speculate that GiAGL1 and GiAG3 could facilitate this process through self-assembly and interactions with the plant cell surface. They are also expected to have other roles, since they are expressed during fungal only growth stages based on RT-PCR (Fig. 3c) and EST data.

## Prediction of post-translational modifications of fungal AGP-like proteins

This is the first report of genes encoding fungal AGP-like proteins, and it is not known if the encoded proteins will be hydroxylated and glycosylated as is characteristic of plant AGPs. The presence of Hyp has been reported in the cell walls of oomycetes such as *Phytopthora* (Novaes-Ledieu et al. 1967; Cassab and Varner 1988), but not in true fungi. Therefore, it is surprising that genes with similarity to the catalytic α-subunit of prolyl 4-hydroxylases are found in the genomes of some, but not all fungi, based on searches of seven completed genomes (Table S2). The absence of Hyp in the cell walls of most fungi suggests that these genes are not functional. We propose that even if prolyl 4-hydroxylases genes are present and functional in *G. intraradices*, many Pro residues in GiAGL1 and GiAGL3 will not be modified because of neighbouring Lys residues (Kieliszewski and Lamport 1994; Johnson et al. 2003; Shimizu et al. 2005). In contrast, GiAGL2 would be readily hydroxylated if prolyl 4-hydroxylase genes are present and active. Until there is experimental evidence showing that GiAGLs are glycosylated, we prefer to use the term "arabinogalactan-protein-like" to reflect their similarity to AGP protein backbones.

Antibodies that detect arabinogalactan moieties of AGPs can be used for immunodetection of fungal tissues to determine if GiAGLs are glycosylated with plant-like carbohydrates. AGP antibodies have been used on mycorrhizal roots of pea (Gollotte et al. 1995) and a variety of other plant species including leek, tobacco and maize (Balestrini et al. 1996). These studies show that there are no AGP epitopes present in the fungal plasma membrane, although in both studies, antibody labelling is detected in the interfacial matrix that could represent soluble AGPs after GPI anchors cleavage (Schultz et al. 1998), and therefore it is not possible to determine whether they are of plant or fungal origin.

Based on knowledge of GPI anchoring in plants (Eisenhaber et al. 2003) and fungi (Eisenhaber et al. 2004), we can be confident that the fungal AGLs will be GPI-anchored. What is uncertain is whether they function at the plasma membrane or are released from the plasma membrane by phospholipase cleavage (Schultz et al. 1998; Pittet and Conzelmann 2007).

## Evolutionary implications of AGL genes from *G. intraradices*

The low similarity between Pro(Hyp)-rich proteins in plants makes it difficult to determine the evolutionary origin of the AGL genes from *G. intraradices.* AGL genes are not found in most fungi (Table S2), and we speculate that the inability to amplify orthologous genes from genomic DNA of *G. versiforme* or from cDNA of *M. truncatula* roots colonised by *G. versiforme* or *Gi. gigantea* is due to diversification of AGLs within the Glomeromycota. Once the genome of *G. intraradices* is available, it will be possible to determine how large the AGL gene family is in *G. intraradices* and identify orthologues from other AM fungi using degenerate primers.

Evidence for differential rates of evolution of Hyp-rich glycoproteins has recently been shown in *Chlamydomonas* where the sexual agglutinins have evolved more rapidly than cell wall Hyp-rich glycoproteins, presumably due to amplification and positive selection of various insertions and deletions (Lee et al. 2007). It is possible that the previously reported *MtAM1* from *G. versiforme*-colonised roots (van Buuren et al. 1999) is the orthologue of *GiAGL3*, since they share 77% nucleotide identity (610/792 nt). Despite earlier data suggesting *MtAM1* is a plant gene (PCR and Southern), it has not yet been found in the *M. truncatula* genome (see Table 1). It is also high in glycine (16.2%), aspartate (4.7%) and lysine (8.4%), unlike

*Arabidopsis* AGPs (Schultz et al. 2002). *MtAM1* and *GiAGL3* are only 52% conserved at the protein level, and the Pro-rich region is less conserved than the ER and GPI signals. Further studies are needed to resolve this uncertainty, but if *MtAM1* is the *G. versiforme* orthologue of *GiAGL3*, the lack of conservation in the mature protein domain supports the potential for the rapid evolution of genes encoding Pro-rich proteins.

The absence of AGL genes from most fungi raises the possibility that the fungal genes arose by horizontal gene transfer from the plant early in the establishment of the AM symbiosis. Horizontal gene transfer is rare in eukaryotes compared to bacteria, but it has been reported and is most common in unicellular eukaryotes (Richardson and Palmer 2007). To determine whether the fungal AGL genes arose from divergent evolution or horizontal gene transfer will be difficult until the genomes of a wide range of Glomeromycota are available.

**Acknowledgements** We gratefully acknowledge Sally Smith and Andrew Smith for helpful discussions and critical comments on the manuscript. We thank Karen Gomez and Stephen Mondo for providing cDNA samples and Damian Drew for assistance with structural modelling using DeepView. Financial assistance was provided from the Australian Research Council (ARC DP0663374), AW Howard Memorial Trust Inc. for a travel grant and National Science Foundation (DBI-0421676).

## References

Armstrong L, Peterson RL (2002) The interface between the arbuscular mycorrhizal fungus *Glomus intraradices* and root cells of *Panax quinquefolius*: a Paris-type mycorrhizal association. Mycologia 94:587–595 doi:10.2307/3761710

Balestrini R, Lanfranco L (2006) Fungal and plant gene expression in arbuscular mycorrhizal symbiosis. Mycorrhiza 16:509–524 doi:10.1007/s00572-006-0069-2

Balestrini R, Hahn MG, Faccio A, Mendgen K, Bonfante P (1996) Differential localization of carbohydrate epitopes in plant cell walls in the presence and absence of arbuscular mycorrhizal fungi. Plant Physiol 111:203–213

Bécard G, Piché Y (1992) Establishment of vesicular–arbuscular mycorrhiza in root organ culture: review and proposed methodology. Methods Microbiol 24:89–108 doi:10.1016/S0580-9517(08)70089-8

Bochicchio B, Tamburro AM (2002) Polyproline II structure in proteins: identification by chiroptical spectroscopies, stability, and functions. Chirality 14:782–792 doi:10.1002/chir.10153

Boisson-Dernier A, Chabaud M, Garcia F, Becard G, Rosenberg C, Barker DG (2001) *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. Mol Plant Microbe Interact 14:695–700 doi:10.1094/MPMI.2001.14.6.695

Bonfante-Fasolo P, Faccio A, Perotto S, Schubert A (1990a) Correlation between chitin distribution and cell wall morphology in the mycorrhizal fungus *Glomus versiforme*. Mycol Res 94:157–165

Bonfante-Fasolo P, Vian B, Perotto S, Faccio A, Knox J (1990b) Cellulose and pectin localization in roots of mycorrhizal *Allium porrum*: labelling continuity between host cell wall and interfacial material. Planta 180:537–547 doi:10.1007/BF02411452

Cassab GI, Varner JE (1988) Cell wall proteins. Annu Rev Plant Physiol Plant Mol Biol 39:321–353 doi:10.1146/annurev.pp.39.060188.001541

Dickson S, Smith FA, Smith SE (2007) Structural differences in arbuscular mycorrhizal symbioses: more than 100 years after Gallaud, where next? Mycorrhiza 17:375–393 doi:10.1007/s00572-007-0130-9

Eisenhaber B, Wildpaner M, Schultz CJ, Borner GHH, Dupree P, Eisenhaber F (2003) Glycosylphosphatidylinositol lipid anchoring of plant proteins. Sensitive prediction from sequence- and genome-wide studies for arabidopsis and rice. Plant Physiol 133:1691–1701 doi:10.1104/pp.103.023580

Eisenhaber B, Schneider G, Wildpaner M, Eisenhaber F (2004) A sensitive predictor for potential GPI lipid modification sites in fungal protein sequences and its application to genome-wide studies for *Aspergillus nidulans*, *Candida albicans*, *Neurospora crassa*, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. J Mol Biol 337:243–253 doi:10.1016/j.jmb.2004.01.025

Gaspar YM, Johnson KL, McKenna JA, Bacic A, Schultz CJ (2001) The complex structures of arabinogalactan-proteins and the journey towards understanding function. Plant Mol Biol 47:161–176 doi:10.1023/A:1010683432529

Gollotte A, Gianinazzi-Pearson V, Gianinazzi S (1995) Immunodetection of infection thread glycoprotein and arabinogalactan-protein in wild type *Pisum sativum* (L.) or an isogenic mycorrhiza-resistant mutant interacting with *Glomus mosseae*. Symbiosis 18:69–85

Grunwald U, Nyamsuren O, Tarnasloukht M, Lapopin L, Becker A, Mann P et al (2004) Identification of mycorrhiza-regulated genes with arbuscule development-related expression profile. Plant Mol Biol 55:553–566 doi:10.1007/s11103-004-1303-y

Harrison MJ (2005) Signaling in the arbuscular mycorrhizal symbiosis. Annu Rev Microbiol 59:19–42 doi:10.1146/annurev.micro.58.030603.123749

Hohnjec N, Henckel K, Bekel T, Gouzy J, Dondrup M, Goesmann A et al (2006) Transcriptional snapshots provide insights into the molecular basis of arbuscular mycorrhiza in the model legume *Medicago truncatula*. Funct Plant Biol 33:737–748 doi:10.1071/FP06079

Johnson KL, Jones BJ, Schultz CJ, Bacic A (2003) Non-enzymic cell wall (glyco)proteins. In: Rose J (ed) The plant cell wall. Blackwell, Oxford, pp 111–154

Kieliszewski MJ, Lamport DTA (1994) Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny. Plant J 5:157–172 doi:10.1046/j.1365-313X.1994.05020157.x

Lee JH, Waffenschmidt S, Small L, Goodenough U (2007) Between-species analysis of short-repeat modules in cell wall and sex-related hydroxyproline-rich glycoproteins of *Chlamydomonas*. Plant Physiol 144:1813–1826 doi:10.1104/pp.107.100891

Liu JY, Blaylock LA, Endre G, Cho J, Town CD, VandenBosch KA et al (2003) Transcript profiling coupled with spatial expression analyses reveals genes involved in distinct developmental stages of an arbuscular mycorrhizal symbiosis. Plant Cell 15:2106–2123 doi:10.1105/tpc.014183

Liu JY, Maldonado-Mendoza I, Lopez-Meyer M, Cheung F, Town CD, Harrison MJ (2007) Arbuscular mycorrhizal symbiosis is accompanied by local and systemic alterations in gene expression and an increase in disease resistance in the shoots. Plant J 50:529–544 doi:10.1111/j.1365-313X.2007.03069.x

Maldonado-Mendoza IE, Dewbre GR, Bell C, Harrison MJ (2004) ESTs from the extra-radical mycelium of *Glomus intraradices*. In: Tikhonovich I, Lugtenberg B, Provorov N (eds) The biology of plant–microbe interactions. International Society for Molecular Plant–Microbe Interactions, Minneapolis, pp 460–464

Nagahashi G, Douds D (2007) Separated components of root exudate and cytosol stimulate different morphologically identifiable types of branching responses by arbuscular mycorrhizal fungi. Mycol Res 111:487–492 doi:10.1016/j.mycres.2007.02.007

Novaes-Ledieu M, Jimenez-Martinez A, Villaneuva JR (1967) Chemical composition of hyphal wall of Phycomycetes. J Gen Microbiol 47:237–245

Peterson RL, Massicotte HB (2004) Exploring structural definitions of mycorrhizas, with emphasis on nutrient-exchange interfaces. Can J Bot 82:1074–1088 doi:10.1139/b04-071

Pittet M, Conzelmann A (2007) Biosynthesis and function of GPI proteins in the yeast *Saccharomyces cerevisiae*. Biochim Biophys Acta–Mol Cell Biol Lipids 1771:405–420

Pujals S, Fernandez-Carneado J, Lopez-Iglesias C, Kogan MJ, Giralt E (2006) Mechanistic aspects of CPP-mediated intracellular drug delivery: relevance of CPP self-assembly. Biochim Biophys Acta–Biomembranes 1758:264–279 doi:10.1016/j.bbamem.2006.01.006

Rauscher S, Baud S, Miao M, Keeley FW, Pomes R (2006) Proline and glycine control protein self-organization into elastomeric or amyloid fibrils. Structure 14:1667–1676 doi:10.1016/j.str.2006.09.008

Richardson AO, Palmer JD (2007) Horizontal gene transfer in plants. J Exp Bot 58:1–9 doi:10.1093/jxb/erl148

Schultz CJ, Coruzzi GM (1995) The aspartate aminotransferase gene family of *Arabidopsis* encodes isoenzymes localized to three distinct subcellular compartments. Plant J 7:61–75 doi:10.1046/j.1365-313X.1995.07010061.x

Schultz CJ, Gilson P, Oxley D, Youl JJ, Bacic A (1998) GPI-anchors on arabinogalactan-proteins: implications for signalling in plants. Trends Plant Sci 3:426–431 doi:10.1016/S1360-1385(98)01328-4

Schultz CJ, Rumsewicz MP, Johnson KL, Jones BJ, Gaspar YM, Bacic A (2002) Using genomic resources to guide research directions. The arabinogalactan-protein gene family as a test case. Plant Physiol 129:1448–1463 doi:10.1104/pp.003459

Seifert GJ, Roberts K (2007) The biology of arabinogalactan-proteins. Annu Rev Plant Biol 58:137–161 doi:10.1146/annurev.arplant.58.032806.103801

Shimizu M, Igasaki T et al (2005) Experimental determination of proline hydroxylation and hydroxyproline arabinogalactosylation motifs in secretory proteins. Plant J 42:877–889 doi:10.1111/j.1365-313X.2005.02419.x

Smith SE, Barker SJ, Zhu YG (2006) Fast moves in arbuscular mycorrhizal symbiotic signalling. Trends Plant Sci 11:369–371 doi:10.1016/j.tplants.2006.06.008

St-Arnaud M, Hamel C, Vimard B, Caron M, Fortin JA (1996) Enhanced hyphal growth and spore production of the arbuscular mycorrhizal fungus *Glomus intraradices* in an in vitro system in the absence of host roots. Mycol Res 100:328–332

Tan L, Leykam JF, Kieliszewski MJ (2003) Glycosylation motifs that direct arabinogalactan addition to arabinogalactan-proteins. Plant Physiol 132:1362–1369 doi:10.1104/pp.103.021766

van Buuren ML, Malonado-Mendoza IE, Trieu AT, Blaylock LA, Harrison MJ (1999) Novel genes induced during an arbuscular mycorrhizal (AM) symbiosis formed between *Medicago truncatula* and *Glomus versiforme*. Mol Plant Microbe Interact 12:171–181 doi:10.1094/MPMI.1999.12.3.171

VandenBosch KA, Stacey G (2003) Summaries of legume genomics projects from around the globe. Community resources for crops and models. Plant Physiol 131:840–865 doi:10.1104/pp.103.020388