Plant and Soil **244:** 129–139, 2002.

Early developmentally regulated genes in the arbuscular mycorrhizal fungus *Glomus mosseae*: identification of GmGIN1, a novel gene with homology to the C-terminus of metazoan hedgehog proteins

Natalia Requena[1,3], Petra Mann[2], Rüdiger Hampp[1] & Philipp Franken[2]
[1]*Botanical Institute, Department of Physiological Ecology of Plants, University of Tübingen, Auf der Morgenstelle 1, D-72076 Germany.* [2]*Max-Planck-Institute for terrestrial Microbiology and Philipps-University of Marburg, Dept. of Microbiology, Karl-von-Frisch-Str., D-35043 Marburg.* [3]*Corresponding author**

Received 21 August 2001. Accepted in revised form 10 March 2002

Key words: arbuscular mycorrhiza, development, gene regulation, *Glomus mosseae*, hedgehog proteins

Abstract

The life cycle of the obligate biotrophic arbuscular mycorrhizal fungi comprises several well-defined developmental stages whose genetic determinants are still unknown. With the aim of understanding the molecular processes governing the early developmental phase of the AM fungal life cycle, a subtractive cDNA library was constructed using a suppressive subtractive hybridization technique. The library contains more than 600 clones with an average size of 500 bp. The isolated cDNAs correspond to genes up-regulated during the early development of the AM fungus *Glomus mosseae* versus genes expressed in extraradical hyphae. The expression of several of the isolated genes was further confirmed by RT-PCR analysis. Among the isolated clones, a novel gene named *GmGIN*1 only expressed during early development in *G. mosseae* was found. The full-length *GmGIN*1 cDNA codes for a protein of 429 amino acids. The most interesting feature of the deduced protein is its two-domain structure with a putative self-splicing activity. The N-terminal domain shares sequence similarity with a novel family of GTP binding proteins while the C-terminus has a striking homology to the C-terminal part of the hedgehog protein family from metazoa. The C-terminal part of hedgehog proteins is known to participate in the covalent modification of the N-terminus by cholesterol, and in the self-splicing activity which renders the active form of the protein with signalling function. We speculate that the N-terminal part of GmGIN1, activated through a similar mechanism to the hedgehog proteins, has GTP-binding activity and participates in the signalling events prior to symbiosis formation.

Abbreviations: AM – arbuscular mycorrhizal; SSH – suppressive subtractive hybridization; sscDNA – single stranded cDNA; RT-PCR – reverse transcription-polymerase chain reaction; hh – hedgehog

Introduction

The mycorrhizal association between the Glomalean fungi and the roots of vascular plants represent for both partners a whole set of morphological and metabolic adaptations to the life in symbiosis. However, it is the fungal partner which depends absolutely on the plant to complete its life cycle. In the absence of it, the arbuscular mycorrhizal spores can only germinate and have a very restricted saprotrophic growth, mainly upon its own carbon expenses (Bago et al., 1999). This growth is very limited in extension but not in time, since the AM fungi have the ability to re-germinate several times. For that, the hyphae expanding out of the spore arrest growth after a certain time, about 2–3 weeks for *G. mosseae*, and suffer apical septation with retrieval of protoplasm towards the spore (Mosse, 1959). This rare mechanism is understood as a survival strategy of the AM fungi under conditions where

* FAX No: 7071-2973083.
E-mail: natalia.requena@uni-tuebingen.de

to find a root partner becomes difficult or impossible (Logi et al., 1998). These spores and growing hyphae have, therefore, a very different role from the extraradical symbiotic hyphae coming out of the root. The first ones are using their stored energy in form of triacylglyerides, glycerol and trehalose to grow enough to find a compatible root, and if not, to arrest growth and suffer a process similar to apoptosis (Bago et al., 1998; Logi et al., 1998). In contrast, the extraradical hyphae have as a main mission to scavenge mineral nutrients from the soil (mainly phosphate) for nursing its symbiotic partner and, eventually, also to colonize other roots. Our goal is to study gene regulation between these two different developmental stages.

The method employed here to identify genes specific from the early developmental stage, the suppressive subtractive hybridization, is meant to isolate transcripts of medium or rare abundance. That is, genes which are not very highly expressed, and that are regulated at that developmental stage. The SSH technique allows then, to identify key regulators of the morphological and metabolic adaptations that the cell suffer to accommodate a specific developmental change. This is one of the first times that this SSH technique has been applied to fungi (Watson et al., 2000). It was originally described for mammalian cells (Diatchenko et al., 1996), and it has been very recently employed for plant cells (Xiong et al., 2001). Other methods, such as the differential RNA display (DDRT-PCR) also described for the isolation of specific and rare transcripts (Lapopin et al., 1999; Martin-Laurent et al., 1997; Requena et al., 1999), or the construction of subtractive libraries (van Buuren et al., 1999) have been previously employed for the isolation of mycorrhizal specific genes. However, DDRT-PCR has been proven to be very time consuming because of the high amount of false positives rendered (Requena et al., 2000), and subtractive libraries allow in principle only the isolation of very highly expressed genes (for review see Franken and Requena, 2001a,b). In the case of fungal genes from arbuscular mycorrhizas, the main handicap is the small amounts of pure fungal material available due to its unculturability. This made SSH our method of choice since it allows starting from amounts of RNA (i.e. tissue) as little as 50 ng.

Materials and methods

Isolation of fungal pre-symbiotic material

Sporocarps from the AM fungus *G. mosseae* (Nicol & Gerd.) Gerdemann & Trappe (BEG12) were purchased from Biorize (France) and surface-sterilized according to Budi et al. (1999). Sporocarps were germinated in water-agar plates as described elsewhere (Requena et al., 1999) and grown in an incubator at 25 °C in the dark. After 12 days, the germinated sporocarps were harvested using fine forceps to recover all hyphae and immediately frozen in liquid nitrogen.

Isolation of fungal symbiotic material

Parsley (*Petroselinum crispum*) plants were inoculated with *G. mosseae* and grown in pots divided into compartments separated by a thin mesh (50 μm) which allows the growth of extraradical hyphae but not of roots (Redecker et al., 1995). The hyphal compartment contained 2 mm glass beads from where the hyphae could be easily recovered. Plants were grown for three months in a phytochamber with controlled conditions of light (100 μE m^{-2} s^{-1}), temperature (20 °C) and photoperiod (14 h light). After that time, extraradical hyphae were isolated from the glass bead compartment, washed several times with chilled distilled water and frozen in liquid nitrogen.

RNA extraction and Subtractive Library Construction

Total RNA was extracted using the RNAeasy kit from Qiagen (Heidelberg) from both the pre-symbiotic material and the extraradical hyphae. RNA quality and quantity was assessed by Northern-blot using a DIG probe containing the ribosomal genes from *G. mosseae* (Franken and Gianinazzi-Pearson, 1996). Equal amounts from both RNA populations (500ng) were used to synthesise cDNA using the SMART protocol from Clontech (Palo Alto, CA). The cDNA populations were then subtracted using the PCR-Select^{TM} cDNA Subtraction kit from Clontech according to manufacturer recommendations, where pre-symbiotic material represented the tester and the hyphal material the driver. Subtracted cDNA fragments were cloned into the TA TOPO vector from Invitrogen (Groningen).

Reverse Northern-blot screening

One hundred clones randomly selected were PCR-amplified using the primers Nested PCR primer 1 and Nested PCR primer 2R (www.clontech.com) that flank the borders of every insert from the subtractive library. PCR was performed in 50 μl using the standard Gibco Taq polymerase, in a Biometra thermal-cycler. Five μl aliquots were separated in a 1.5% agarose gel and Southern blotted in parallel into two nylon membranes according to Sambrook et al. (1989). A radioactive probe was prepared using as template SMART cDNA from each population (pre-symbiotic mycelium and extraradical hyphae) by incorporation of $\alpha-^{32}$P dATP in a PCR reaction using as primers the SMART PCR primer from the SMART kit. The membranes were then hybridized using the standard DIG buffer (Boehringer) at 65 °C overnight. Membranes were then washed twice using 2X SSC; 0.1% SDS buffer at room temperature, and one time with 0.5XSSC; 0.1% SDS buffer at 65 °C. Membranes were then exposed for 24–72 h to an X-ray film.

Sequencing and homology searches

Clones which were found to be either up or down regulated in the Reverse Northern-blots were sequenced and their sequence compared to the non-redundant NCBI database using the programme BlastX 2.0.11 (Altschul et al., 1997). The parameters used for the sequence comparison were an expected value of 10, Word size 3, using the Blosum62 matrix with Gap cost: Existence 11, Extension 1. Only sequences showing an E value lower than 10 and with an overall homology higher than 40% were considered as significant homologies and are listed in Table 1. Sequences listed in Table 1 were submitted to the EMBL database and Accession numbers are given.

RT-PCR expression analyses

To confirm the expression pattern observed in the Reverse Northern-blot, the expression of ten of the sequenced cDNA fragments was further studied by RT-PCR. Single stranded cDNA was synthesised from 50 ng of total RNA using an oligo (dT) primer. cDNA was synthesised from each population using the MMLV reverse transcriptase from Promega at 37 °C for 1 h. Specific primers designed for each of the genes analyzed were used (Table 2). One microliter of a dilution 1:5 of the sscDNA was used as template for a 30 cycles PCR reaction. The analysis was repeated twice.

Isolation of the full length cDNA of the gene GmGIN1

A modified RACE (rapid amplification of cDNA ends) technique was used to amplify the 5′ and 3′ ends of the cDNA encoding the gene *GmGIN*1. A nested PCR reaction was performed on the SMART cDNA using two pairs of primers based on the sequence of the cDNA fragment in combination with either the PCR primer (SMART kit) for the 5′ end or the CDS primer (SMART kit) for the 3′ end. The PCR reactions were performed using the proof-reading enzyme Advantage 2 (Clontech). The resulting fragments were cloned into the pCR2.1 TOPO vector (Invitrogen) and sequenced (MWG Biotech, Germany).

PCR amplification of GmGIN1 in genomic DNA

Genomic DNA from *Glomus mosseae* was isolated from germinated sporocarps (100) with the Qiagen DNeasy Plant Mini kit. Two PCR reactions were performed. In the first one, the primers used (Gm79F2/R2) amplified the region encoding the C-terminus of *GmGIN*1. In the second one, a combination of primers (Gm79F1/R2) was used that amplify a larger region that covers part of the N-terminus as well as part of the C-terminus of GmGIN1. Both PCRs used as template150pg of genomic DNA in a 25 μl reaction. Primers are listed in Table 2.

Expression of GmGIN1 during different stages of the life cycle

Germinated sporocarps, as well as extraradical hyphae and colonized roots were used as template to study the expression of *GmGIN*1. The cDNA was synthesised with hexamers primers using the Superscript II reverse transcriptase from GIBCO (Life Technologies, UK) according to manufacturer recommendations. Equal amounts of total RNA (70 ng) were used for cDNA synthesis except for the colonized root were five times more RNA belonging to plant and fungus together were employed. cDNA amounts were calibrated by PCR using the ribosomal primers VANS1 and VA-GLO (Simon et al., 1992, 1993). Equal fungal cDNA amounts were then used for the expression analysis.

Table 1. List of isolated cDNA clones from *G. mosseae* expressed during pre-symbiosis with significant homologies to proteins from the NCBI database

Clone	Acc. number	Size (bp)	Position	Definition	Organism	Acc. number	E value	Homol.
1	AJ315702	374	polyA	spindle pole body component Alp6	*Schizosaccharomyces pombe*	BAA94097	0.002	41%
9	AJ315703	265	?	TrihydrophobinCpth	*Claviceps fusiformis*	CAB61236	0.66	43%
11	AJ315704	455	3′ end	proteasome subunit, beta type, 2	*Rattus norvegicus*	NP_058980	1e-15	71%
14	AJ315705	807	5′	F-box protein Fb12	*Homo sapiens*	AAF04510	1.5	45%
16	AJ315706	194	5′	60S ribosomal protein L15	*S. pombe*	T41421	6.6	93%
18	AJ315707	546	mid.	P-type H+ATPase	*Pneumocystis carinii*	AAB06958	1e-29	59%
69	AJ315708	427	mid.	Sec14	*S. pombe*	Q10137	9e-39	72%
79	AJ315709	1132	mid.	N-terminus: AIG1 protein	*Arabidopsis thaliana*	AAG52215	1e-12	56
				C-terminus:Desert hedghegog	*Xenopus laevis*	Q91611	2e-12%	51%
80	AJ315710	362	mid.	28S ribosomal RNA gene	*Lobaria pulmonaria*	AF183934	3e-22	94%
85	AJ315711	465	mid.	DNAJ tetratricopeptide	*S. pombe*	CAC05244	7e-21	54%
89	AJ315712	293	mid.	cytosolic leucine aminopeptidase	*Rickettsia prowazekii*	P27888	2e-17	67%
92	AJ315713	348	3′	purine nucleotide binding protein Fet5	*S. pombe*	T38803	1e-09	78%
101	AJ315714	425	5′	hypothetical protein YNR020p	*Saccharomyces cerevisiae*	NP_014417	2e-06	55%
111	AJ315715	250	3′	CG3060 gene product	*Drosophila melanogaster*	AAF47123	1e-07	60%
118	AJ315716	425	midd.	hypothetical protein T08B2.5	*Caenorhabditis elegans*	T28754	0.046	57%
135	AJ315717	500	poly A	Glucose repression regulatory protein RGR1	*S. cerevisiae*	P19263	2.4	50%
142	AJ315718	308	poly A	Polyubiquitin	*Neurospora crassa*	UQNC	3e-04	60%
149	AJ315719	405	3′	alpha NAC	*D. melanogaster*	AAB97513	4e-38	78%
512	AJ315720	942	midd.	Inositol polyphosphate 5-phosphatase	*H. sapiens*	Q01968	8e-53	57%
513	AJ315721	586	midd.	ubiquitin carboxyl-terminal hydrolase	*A. thaliana*	AAK25908	0.005	53%
519	AJ315722	507	5′	phosphatidylglycerol/phosphatidylinositol of transfer protein	*Aspergillus oryzae*	AAD16095	2.0	48%
520	AJ315723	342	5′	Vacuolar ATP synthase subunit E	*N. crassa*	Q01278	7e-23	82%
1000	AJ315724	1045	5′	CGI-15 protein	*H. sapiens*	XP_009673	6e-27	56%
				similar to phosphate/phosphoenol piruvate translocator				
1001	AJ315725	594	5′	CDC48	*Dyctiostelium discoideum*	U83085	4e-51	78%
1002	AJ315726	292	5′	Putative 3-Ketoacyl-CoA Thiolase	*Aspergillus oryzae*	Q9HG46	6e-16	70%
1003	AJ315727	578	midd.	Heat shock protein 70	*P. carinii*	AAD00455	4e-57	91%
1004	AJ315728	791	5′	Splicing factor, arginine/serine-rich 4	*H. sapiens*	XP_001807	7e-21	75%

Results

Establishment of a suppressive subtractive library

SSH was employed as described in the 'Material and methods' section to obtain a library of cDNA clones up-regulated during early development. A total of 625 colonies containing cDNA fragments from genes expressed during early development in *G. mosseae* were isolated. From those, about 9% of the clones contained either no insert or a very small insert to be seen on an agarose gel. One hundred clones were PCR amplified using primers Nested PCR primer 1 and Nested PCR primer 2R to analyse their expression by reverse-northern blot. Most of the clones gave a single amplification signal as expected but some of them showed more than one band. This can be the result of artefacts during cloning in which more than one cDNA fragment are ligated into the vector (Figure 1). From those 100 clones, 70 were sequence analysed; none of the sequences was redundant, and only two or three of them could potentially belong to different fragments from the same gene. Only one ribosomal sequence corresponding to the large ribosomal subunit and one 60S ribosomal protein were found. The average size of the library clones is around 500 bp.Twenty seven percent of the clones corresponded to the poly A

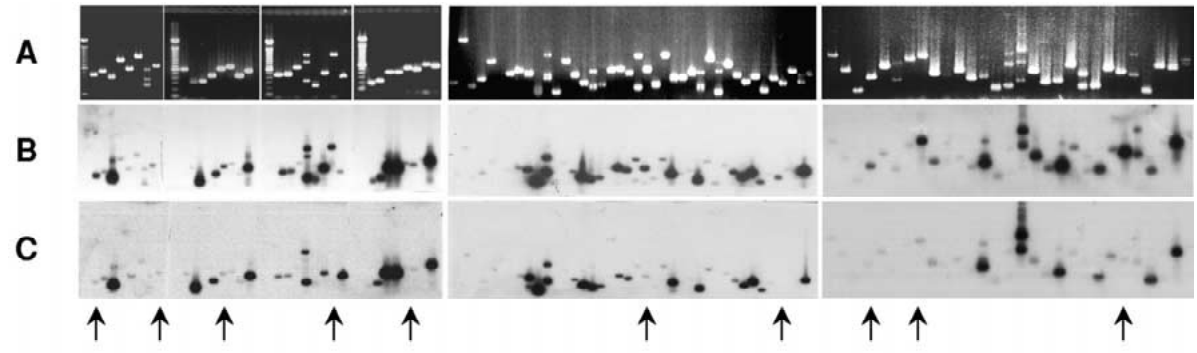

Figure 1. Reverse-northern blot screening for identification of regulated genes between pre-symbiotic and symbiotic hyphae. Panel A shows the PCR amplification of 100 clones from the cDNA library using the primers Nested1 and Nested2R resolved in an agarose gel and stained with ethidium bromide. Panel B and C show the differential hybridization of the nylon membranes, where the clones were blotted, with radioactive labelled cDNA synthesised either from pre-symbiotic material (B) or from extraradical hyphae (C). Clones further analysed by RT-PCR are indicated by arrows.

Table 2. List of the primers used in this paper

Primer name	Sequence
Gm18F3	ccagtcaagaaagttcccaag
Gm18R1	cacgttcagctgctgatatcg
Gm79F2	ggagactacgtatgttgtgg
Gm79R2	aacaagagcagagaacctcg
Gm79F1	gatgtctcggattcagctgc
Gm85F1	gtcaggaagcattgcgttgc
Gm85R1	catatcactcaatgcctccg
Gm1000F1	ggtagtaacgatagttcaccacc
Gm1000R1	gcggatcatctgttcatacttcc
Gm111F1	ttggaatttgaagtctcgcc
Gm11R1	tcttcatttgatacgaaccc
Gm520F1	atgtcgaatctacgtgc
Gm520R1	agaagcagcttgaatgg
Gm1001F1	ctacataccttttacctgatgagc
Gm1001R1	ctcttgtgatttgtggaactgg
Gm1003F1	agcgaccaagaagttcaatccg
Gm1003R1	gacatcaaaaactccaccaccc
Gm69F1	cgtgaagaattcaaggaaggagg
Gm69R1	caggcaggtaatctgacgtcagc
Gm512F1	atccga ttgtgtcggtgtcg
Gm512R1	gatccaatcatccctcgcac

tail while only 15% to the 5′ end of the corresponding genes. Only 27 of the sequenced clones (38%) had a significant homology (Table 1) and from those only four clones were located at the poly A tail. Most of the sequences had homology to fungal proteins (60%) while the rest showed homology to proteins from different eukaryotes, and only one showed significant homology to a bacterial sequence. The GC content of the analysed sequences was always lower than 50% which seems to be characteristic of arbuscular mycorrhizal fungi (Hosny et al., 1997). In several cases where the cDNA clone corresponded to the 5′ region of the gene, a large poly T region could be observed in the 5′ untranslated region (data not shown).

Expression analyses

The Reverse-Northern blot experiment (Figure 1) showed that as expected, the membranes hybridized with the pre-symbiotic labelled cDNA gave a more intense hybridization signal than those hybridized with labelled hyphal cDNA. There was one exception, clone Gm80 (AJ315710), which belonged to the ribosomal large subunit gene. Some clones, such as clone Gm79 (AJ315709), with similarity to GTPases and hedgehog proteins, showed high level of regulation in this screening with a high expression during pre-symbiosis in contrast to no signal in extraradical mycelium.

The RT-PCR analysis of the 10 analysed clones revealed that most of them were indeed up-regulated during pre-symbiosis versus extraradical hyphae. Half of the clones showed at least more than three times up-regulation while the other half were only slightly up-regulated. Only one of the analysed clones (Gm69) showed no changes in gene expression between pre-symbiosis and extraradical hyphae, while in another one (Gm512) the level of expression was higher in hyphae, than in spores.

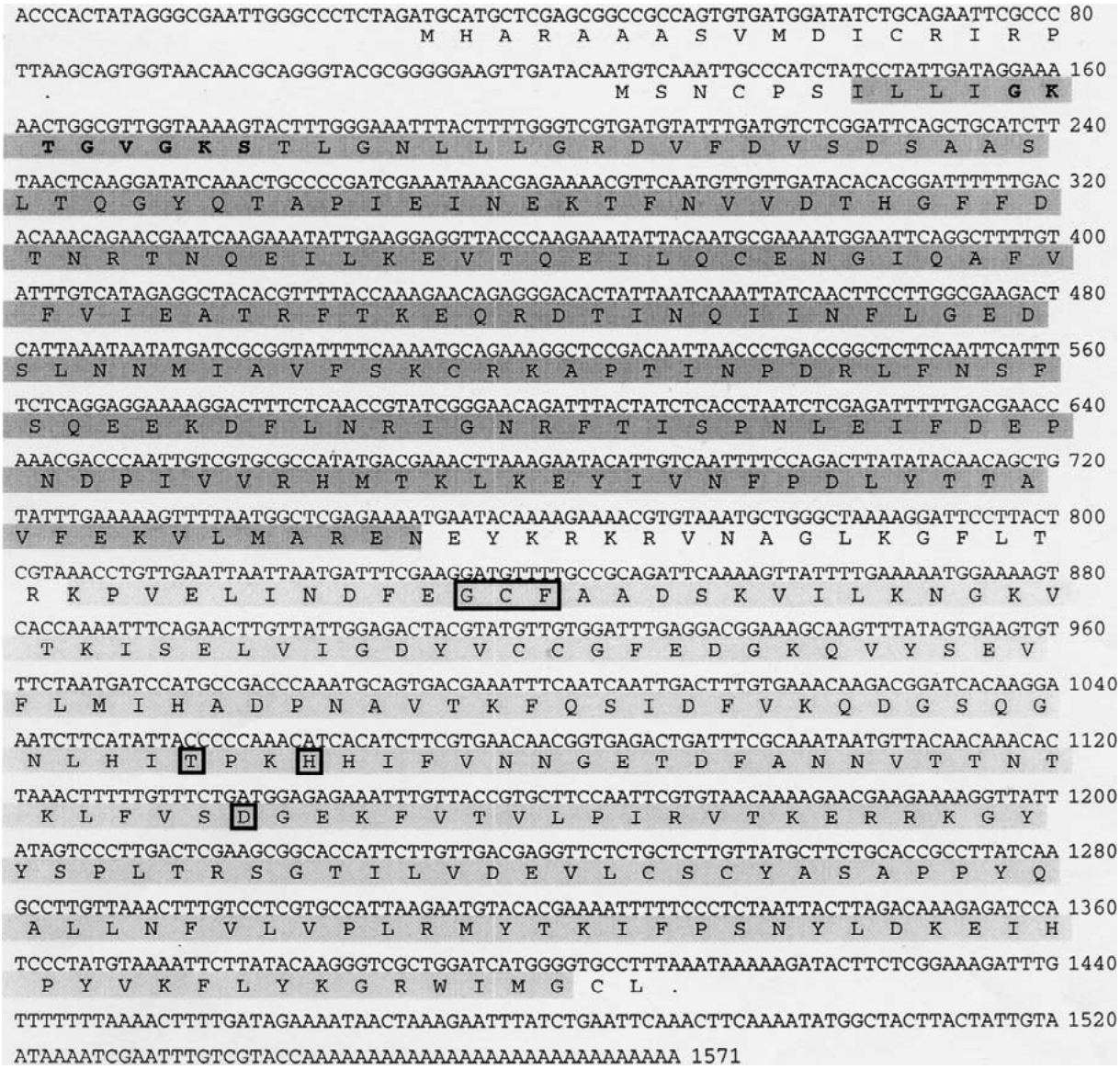

Figure 2. Full length cDNA from *GmGIN*1 showing the large open reading frame encoding the 429 aa GmGIN1 protein. The GTP CDC motif (PF 00735) which extends through most of the amino terminus of the protein is highlighted in dark grey. Inside of that region the nucleotide binding site (ATP/GTP P-loop, PS00017) is indicated in bold. The carboxy-terminus of the protein contains the Hint domain (PF01079) characteristic of hedgehog proteins (highlighted in light grey). Conserved amino acids involved in the autoprocessing of the protein or in cholesterol binding are boxed.

Analysis of GmGIN1

The amplification of 5′ and 3′ cDNA ends from the clone Gm79 rendered a 1571bp fragment with a large open reading frame (ORF) of 1287 bp and the corresponding gene was called *GmGIN1* (Figure 2). The deduced protein, GmGIN1, contains 428 aa and has a calculated molecular mass of 48 kDa. A small ORF (17 aa) was found 96 bp upstream of the first codon. A profile search of the deduced protein using the Pfam-A and Prosite data banks (www.isrec.isb-sib.ch/cgi-bin/PFSCAN) showed three significant matches. The first 209 aa matched to the GTP CDC motif (PF 00735) characteristic of proteins involved in cell division, shadowed in dark grey in Figure 2, with an ATP/GTP binding domain (PS00017) located between amino acids 11–18 (in bold in Figure

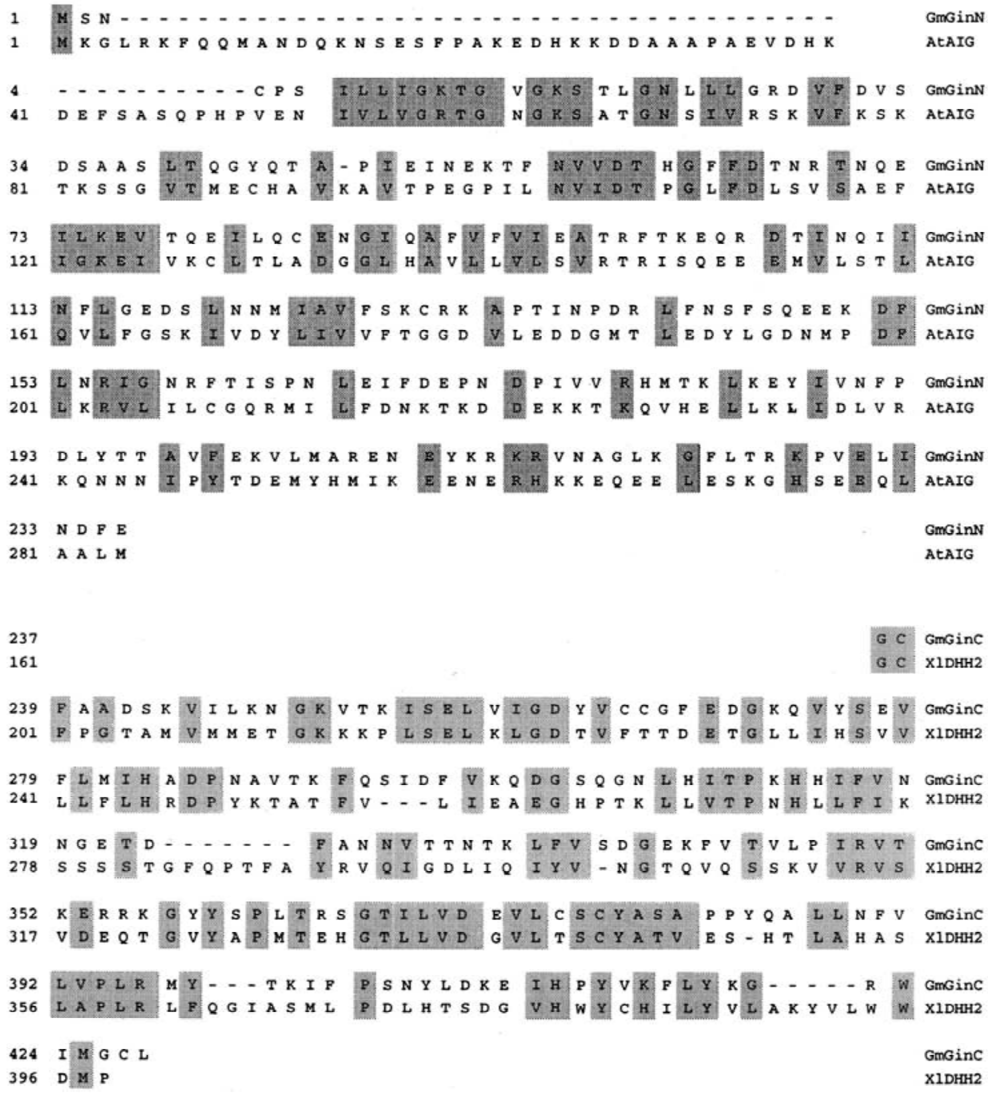

Figure 3. Alignment of the amino and carboxy terminus from the predicted GmGIN1 protein with their closest homologues, *Arabidopsis thaliana* AIG1 (At AIG, AAG52215) and *Xenopus laevis* Desert hedgehog (Xl DHH2, Acc. Q91611) respectively. Conserved amino acids are highlighted.

2). A Hint module (PF01079) characteristic of hedgehog proteins and inteins (proteins that suffer splicing as post-translational modification) was found between residues 227 and 426 (in light grey in Figure 2).

The similarity search of the large protein in the NCBI database using the programme BlastP showed that the C-terminal part of the gene has high homology (50%) to the C-terminal domain of the hedgehog protein family from vertebrates. This conserved domain is not found outside the metazoa and indeed, no significant homology to any fungal protein was found when the GmGIN C-terminus was blasted against fungal databases. Several conserved amino acids essential for the autoproteolysis activity catalyzed by the C-terminus of hedgehog proteins were found. Thus, a conserved cystein residue (Cys-236 in GmGIN1) in the conserved tripeptide GlyCysPhe (GCF), is known to be essential for the initial nucleophilic attack as well as for cholesterol binding, while Thr-310 and His-313 are essential for auto-cleavage (boxed in Figure 2). An aspartic residue involved in cholesterol binding was also found (Asp-285).

The N-terminal part of GmGIN1 showed significant homology (46%) to several proteins with GTPase

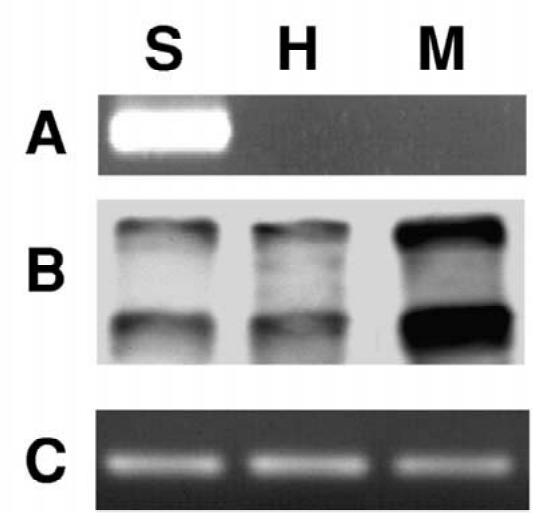

Figure 4. Differential expression of GmGIN1 during distinct developmental stages of the life cycle of *G. mosseae*, S (early development), H (extraradical hyphae), and M (mycorrhizal roots). Panel A shows the expression of the gene studied by RT-PCR and analysed on a 1.5% agarose gel and stained with ethidium bromide. Panel B shows the northern blot calibration with the amount of RNA used from each sample to synthesise the cDNA. Panel C shows the cDNA calibration from each sample by PCR using the primers VANS1 and VAGLO specific for *G. mosseae* and analysed on agarose gel and ethidium bromide staining. One microliter of a 1:10 000 dilution of the synthesized cDNA was used as template in a 28, 30 and 32 cycles PCR. Panel C shows the result after 30 cycles.

activity, and particularly to a novel family of GTP-binding proteins found in plants and more recently in animals. Weak similarities were found between the amino terminus of GmGIN1 and the septin family of fungal proteins probably due to the presence of a CDC type GTP binding domain. Alignments of the N- and the C-terminus from GmGIN1 with their respective putative homologues are presented in Figure 3. Because GmGIN1 showed similarity to different proteins at both ends, we performed a PCR analysis on genomic DNA to exclude the possibility of a PCR artefact during the construction of the library. The PCR analysis performed on genomic DNA confirmed that indeed both ends of the cDNA are part of a single gene.

The analysis of *GmGIN*1 expression was studied at different developmental stages of the fungal life cycle. Thus, while the gene is highly expressed during presymbiosis it seems to be completely shut down during symbiosis, both at the intraradical structures and in the external mycelium (Figure 4).

Discussion

Early developmental genes in the life cycle of arbuscular mycorrhizal fungi must encode proteins responsible for spore germination as well as for the pre-symbiotic hyphal growth in search of a host plant. Some of those should be in charge of the plant recognition and further signaling events that lead to the formation of the symbiosis. In contrast, other genes have to regulate the growth arrest and cytoplasm retraction that takes place in the absence of signals from a compatible host root. Our aim when constructing the library described in this paper was to isolated genes related to the signaling events taking place in the fungus prior to symbiosis formation.

The suppressive-subtractive library constructed in this paper contains 625 clones, about the same number of other SSH libraries described (Xiong et al., 2001). The average size of the clones produced, 500 bp, were demonstrated to be appropriate to find significant homologies in the databases with a good degree of certainty. The reverse-northern blot screening served as a pre-selection for sequencing of highly regulated genes. However, this method underestimates the number of regulated genes because many of the clones showed very weak or no signal even after several days of exposure. They probably belong to genes expressed at very low levels that might be then underscored. A detailed analysis by RT-PCR for those clones would probably disclose their expression pattern. The sequence similarity analysis of the sequenced clones to known proteins from the NCBI database revealed that only a small percentage corresponds to highly conserved proteins whose identity could be predicted. The majority of the sequences, however, belonged to evolutionarily not so highly conserved proteins or to untranslated cDNA regions. The RT-PCR analyses of some of the sequenced clones confirmed their expression pattern and showed that indeed the SSH method allowed identification of specifically regulated genes during early development.

Among those regulated genes, *GmGIN*1 appears to be a novel gene encoding a protein with two different domains. The amino-terminus of GmGIN1, GmGIN1-N has homology with a series of proteins all containing a nucleotide-binding site. In fact, a characteristic ATP/GTP binding site A (P-loop) motif was found between amino acids 11 and 18. A significant match with the GTP CDC cell division motif was also found between amino acids 7 and 209 which constitute most of the N-terminal part of the protein. This motif is

present in proteins related to control of cell cycle, i.e. septins, and they all contain a GTP binding site. However, the similarity search in the databases showed the highest similarity with a new family of GTP-binding proteins first described in plants, and now recently also in mammals, which are pathogen-induced and could be related to the control of an apoptosis response (Poirier et al., 1999; Reuber and Ausubel, 1996). It seems that GmGIN1-N could be a new member of this class of proteins and probably also possess GTPase activity.

The C-terminal domain of GmGIN1, GmGIN1-C has a striking similarity to the C-terminal domain of hedgehog (hh) proteins. The carboxy-terminus domain from hh proteins functions as a cholesterol transferase. It catalyzes the autoprocessing of the protein that renders an amino-terminus covalently modified by a cholesterol moiety (Hall et al., 1997; Porter et al., 1996a, b). Such modification produces the active form of the protein as well as determining its tissue specific spatial localization at the cell membrane (Ingham, 2000; Mann and Beachy, 2000; Porter et al., 1996a). The mechanism by which such a process takes place is a nucleophilic attack by the thiol group of a very conserved cystein (in our case Cys-236) to the carbonyl group of the preceding amino acid residue (a glycine) to form a thioester linkage in place of the peptide bond. In a second step, the thioester suffers also nucleophilic attack from the 3β hydroxl group of a cholesterol molecule, releasing the N-terminal part of the protein covalently attached to cholesterol. The attachment of cholesterol increases dramatically the hydrophobicity of the protein that is then targeted to the cell membrane (Porter et al., 1996a). The mechanism has similarities to the protein splicing mechanism of inteins (Paulus, 2000), and therefore proteins having such a conserved hh carboxy-terminus are recognized in the database to contain the so-called Hint (*H*edgehog *int*ein) motif.

The whole autoprocessing mechanism in hh proteins seems to be independent of the N-terminus (Hall et al., 1997; Porter et al., 1996a, b). Interestingly, a number of unrelated protein families from the worm *Caenorhabditis elegans* have been recently found to possess similar carboxy-terminal regions albeit a totally non-related amino terminus. These proteins have been found to be cleaved in a similar manner to hedgehog proteins (Burglin, 1996; Porter et al., 1996a), although the participation of sterols other than cholesterol is presumed (Aspöck et al., 1999). Only members from the metazoa have been found, so far, to belong to this family of hh C-terminus homologues. GmGIN1 represents, therefore, the first member of this family from lower eukaryotes. A detailed analysis of active sites in the C-terminal domain of hedgehog proteins showed that, besides the above mentioned cystein, two other conserved downstream amino acids, histidine and threonine had an essential role in the autoprocessing (Hall et al., 1997). When mutated, the formation of the internal thioester and the autocleaving activity is severely hindered. In GmGIN1 they are also conserved and correspond to His-310 and Thr-313, indicating a possible similar mechanism of action. The question is whether cholesterol catalyses also the second nucleophilic attack in *G. mosseae*. An important amino acid, an aspartic residue about 50 amino acids downstream of the nucleophilic cystein was found to be dispensable for the first nucleophilic attack but required for the cholesterol transfer (Hall et al., 1996). This residue found in hedgehog proteins is not apparent in inteins sequences, where cholesterol is not the responsible of the processing activity. In GmGIN1, an aspartic residue (Asp-285) is located 47 amino acids downstream the Cys-236. It raises the possibility that in *G. mosseae* cholesterol, or another sterol as in the *C. elegans* proteins, could be responsible for the nucleophilic attack to the thioester formed between the cystein and the preceding glycine. AM fungi are known to contain large amounts of lipids that serve as the main carbon source during pre-symbiotic growth. It would not be surprising that they would also be involved in the signaling processes taking place at this stage.

Our current hyphothesis is that GmGIN1 suffers autoprocessing in a similar manner to hh proteins (see Figure 5). A sterol transfer activity located in the carboxy-terminus would provoke the cleavage of the protein between the amino acids 236 and 237 releasing an amino terminal domain attached to a cholesterol or another sterol moiety. This would cause the activation and proper localization of the N-terminal part of the protein with GTPase activity at the plasma membrane. It remains to be determined which precise function this GTPase would have. Because the expression data show that the gene is mostly expressed in the pre-symbiotic stage prior to contact with the plant, this GTPase activity could be involved in the signaling cascade controlling growth arrest and further programmed cell death in the absence of a signal from the host plant.

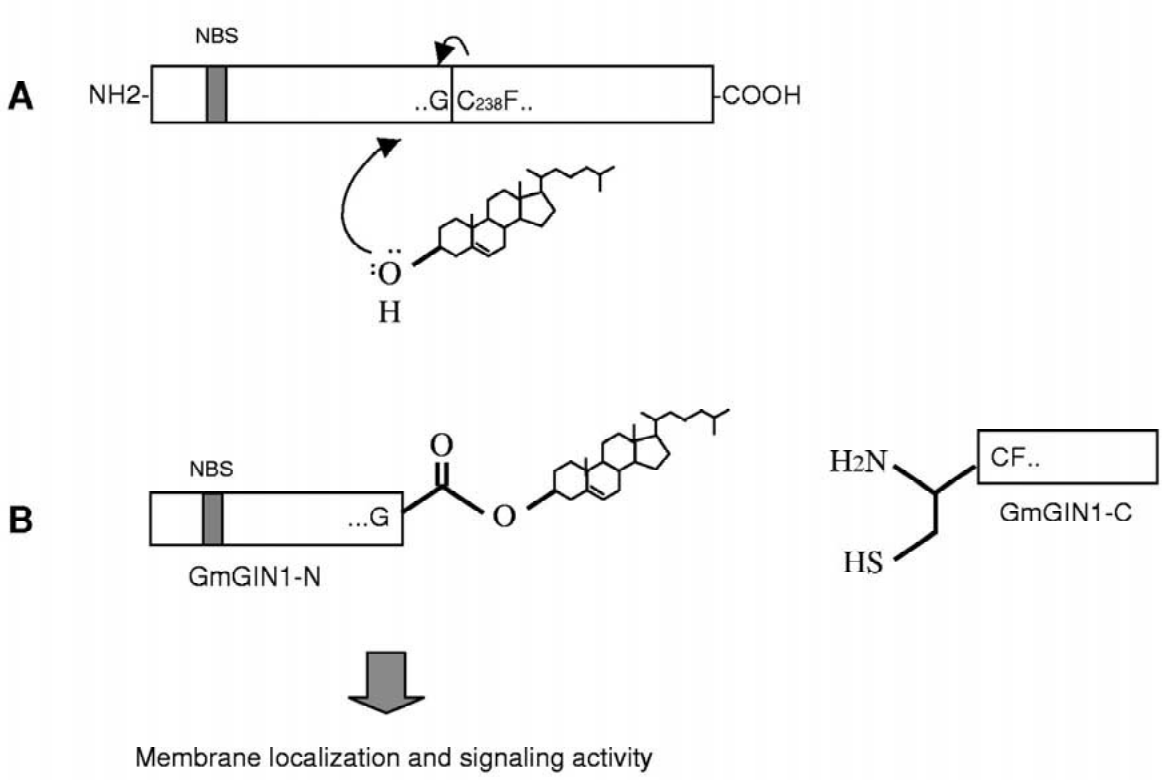

Figure 5. Diagram of the proposed autoprocessing reaction of GmGIN1 according to the model proposed for hedgehog proteins (Hall et al., 1997). A shows the protein structure of GmGIN1 with the nucleotide binding site (NBS) in the N-terminus and the GCF tripeptide containing the cysteine-238 in the middle of the protein. The Cys-238 initiates the nuclephilic attack on the carbonyl carbon to the precedent amino acid residue. This attack forms an intermediate thioester linkage between Gly-237 and Cys-238 which is resolved with the second nucleophilic attack performed by the electrons from the 3β hydroxyl group of a cholesterol molecule. As a result, the protein is spliced in two and the N-terminus results covalently attached to a cholesterol moiety (B). This modification increases considerably the hydrophobicity of the N-terminal part of the protein that may then localize to the membrane and exert a signaling function through its GTPase activity.

Acknowledgements

We thank A. Ocón-Garrido, V. Karidaki and B. Achatz for their help with the expression studies. Special thank to Drs. R. Fischer and M. Takka for critical reading of the manuscript and helpful discussions.

References

Altschul, S F, Madden T L, Schäffer A A, Zhang J, Zhang Z, Miller W and Lipman D J 1997 Gapped BLAST and PSI BLAST: A new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Aspöck Gq, Kagoshima H, Niklaus G and Bürglin T R 1999 *Caenorhabditis elegans* has scores of hedgehog-related genes: Sequence and expression analysis. Genome Research 9, 909–923.

Bago B, Zipfel W, Williams R M, Chamberland H, Lafontaine J G, Webb W W and Piché Y 1998 In vivo studies of the nuclear behaviour of the arbuscular mycorrhizal fungus *Gigaspora rosea* grown under axenic conditions. Protoplasma 203, 1–15.

Bago B, Pfeffer P E, Douds D D Jr, Brouillette J, Becard G and Shachar-Hill Y 1999 Carbon metabolism in spores of the arbuscular mycorrhizal fungus *Glomus intraradices* as revealed by nuclear magnetic resonance spectroscopy. Plant Physiol 121, 263–72.

Budi S W, Blal B and Gianinazzi S 1999 Surface-sterilization of *Glomus mosseae* sporocarps for studying endomycorrhization *in vitro*. Mycorrhiza 9, 65–68.

Burglin T R 1996 Warthog and Groundhog, novel families related to Hedgehog. Curr. Biol. 15, 2294–2303.

Diatchenko L, Lau Y F, Campbell A P, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov E D and Siebert P D 1996 Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. USA 93, 6025–6030.

Franken P and Gianinazzi-Pearson V 1996 Phage cloning of ribosomal RNA genes from the arbuscular mycorrhizal fungi *Glomus mosseae* and *Scutellospora castanea*. Mycorrhiza 6, 167–173.

Franken P and Requena N 2001a Molecular approaches to arbuscular mycorrhiza functioning. *In* The Mycota IX. Fungal Associations. Ed. B Hock. pp 19–28. Springer-Verlag, Berlin.

Franken P and Requena N 2001b Analysis of gene expression in arbuscular mycorrhizas: New approaches and challenges. New Phytol. 150, 517–523.

Hall T M, Porter J A, Young K E, Koonin E V, Beachy P A and Leathy D J 1997 Crystal structure of a hedgehog autoprocessing domain: homology between hedgehog and self-splicing proteins. Cell 91, 85–97.

Hosny M, Pais de Barros J P, Gianinazzi-Pearson V and Dulieu H 1997 Base composition of DNA from glomalean fungi: high amounts of methylated cytosine. Fungal Genet. Biol. 22, 103–11.

Ingham P W 2000 How cholesterol modulates the signal. Curr. Biol. 10, R180–R183.

Lapopin L, Gianinazzi-Pearson V and Franken P 1999 Comparative differential RNA display analysis of arbuscular mycorrhiza in *Pisum sativum* wild type and a mutant defective in late stage development. Plant Mol. Biol. 41, 669–77.

Logi C, Sbrana C and Giovannetti M 1998 Cellular events involved in survival of individual arbuscular mycorrhizal symbionts growing in the absence of the host. Appl Environ Microbiol 64, 3473–3479.

Mann R K and Beachy P A 2000 Cholesterol modification of proteins. Biochim. Biophys. Acta1529, 188–202.

Martin-Laurent F, van Tuinen D, Dumas-Gaudot E, Gianinazzi-Pearson V, Gianinazzi S and Franken P 1997 Differential display analysis of RNA accumulation in arbuscular mycorrhiza of pea and isolation of a novel symbiosis-regulated plant gene. Mol. Gen. Genet. 256, 37–44.

Mosse B 1959 The regular germination of resting spores and some observations on the growth requirements of an *Endogone* sp. causing vesicular-arbuscular mycorrhiza. Trans. Br. Mycol. Soc. 42, 273–286.

Paulus H 2000 Protein splicing and related forms of protein autoprocessing. Ann. Rev. Biochem. 69, 447–496.

Poirier G M C, Anderson G, Huvar A, Wagaman P C, Shuttleworth J, Jenkinson E, Jackson M R, Peterson P A and Erlander M G 1999 Immune-associated nucleotide-1 (IAN-1) is a thymic selection marker and defines a novel gene family conserved in plants. J. Immunol. 163, 4960–4969.

Porter J A, Ekker S C, Park W-J, von Kessler D P, Young K E, Chen C H, Ma Y, Woods A S, Cotter R J, Koonin E V and Beachy P A 1996a Hedgehog patterning activity: Role of a lipophilic modification mediated by the carboxy-terminal autoprocessing domain. Cell 86, 21–34.

Porter J A, Young K E and Beachy P A 1996b Cholesterol modification of Hedgehog signaling proteins in animal development. Science 274, 255–297.

Redecker D, Thierfelder H and Werner D 1995 A new cultivation system for arbuscular mycorrhizal fungi on glass beads. Angewandte Botanik 69, 189–191.

Requena N, Fuller P and Franken P 1999 Molecular characterization of GmFOX2, an evolutionarily highly conserved gene from the mycorrhizal fungus *Glomus mosseae*, down-regulated during interaction with rhizobacteria. Mol. Plant-Microbe Interact.12, 934–42.

Requena N, Mann P and Franken P 2000 A homologue of the cell cycle checkpoint TOR2 from *Saccharomyces cerevisiae* exists in the arbuscular mycorrhizal fungus *Glomus mosseae*. Protoplasma 212, 89–98.

Reuber T L and Ausubel F M 1996 Isolation of *Arabidopsis* genes that differentiate between resistance responses mediated by the RPS2 and RPM1 disease resistance genes. Plant Cell 8, 241–249.

Sambrook J, Fritsch E F and Maniatis T A 1989 Molecular cloning: a laboratory manual. 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Simon L, Lalonde M and Bruns T D 1992 Specific amplification of 18S fungal ribosomal genes from vesicular-arbuscular endomycorrhizal fungi colonizing roots. Appl. Environ. Microbiol. 58, 291–295.

Simon L, Lévesque R C and Lalonde M 1993 Identification of endomycorrhizal fungi colonizing roots by fluorescent single-strand conformation polymorphism-polymerase chain reaction. Appl. Environ. Microbiol. 59, 4211–4215.

van Buuren M L, Maldonado-Mendoza I E, Trieu A T, Blaylock L A and Harrison M J 1999 Novel genes induced during an arbuscular mycorrhizal (AM) symbiosis formed between *Medicago truncatula* and *Glomus versiforme*. Mol. Plant Microbe-Interact. 123, 171–81.

Watson, A J, Worley J, Elliott R M, Jeenes D J and Archer D B 2000 Cloning stress-induced genes from *Aspergillus niger* using polymerase chain reaction-augmented subtractive hybridization. Anal Biochem 277, 162–165.

Xiong L, Lee M-W, Qi M and Yang Y 2001 Identification of defense-related rice genes by suppression subtractive hybridization and differential screening. Mol. Plant Microb. Interact. 14, 685–692.