Comparison of gene expression in trap cells and vegetative hyphae of the nematophagous fungus Monacrosporium haptotylum

Ahrén, Dag; Tholander, Margareta; Fekete, Csaba; Rajashekar, Balaji; Friman, Eva; Johansson, Tomas; Tunlid, Anders

Published in:
Microbiology

DOI:
10.1099/mic.0.27485-0

2005

Microbiology (2005), **151**, 789–803 DOI 10.1099/mic.0.27485-0

Comparison of gene expression in trap cells and vegetative hyphae of the nematophagous fungus *Monacrosporium haptotylum*

Dag Ahrén,†‡ Margareta Tholander,‡ Csaba Fekete,‡ Balaji Rajashekar, Eva Friman, Tomas Johansson and Anders Tunlid

Department of Microbial Ecology, Lund University, Ecology Building, SE-223 62 Lund, Sweden

Correspondence
Anders Tunlid
Anders.Tunlid@mbioekol.lu.se

Received 14 July 2004
Revised 8 November 2004
Accepted 10 November 2004

Nematode-trapping fungi enter the parasitic stage by developing specific morphological structures called traps. The global patterns of gene expression in traps and mycelium of the fungus *Monacrosporium haptotylum* were compared. The trap of this fungus is a unicellular spherical structure called the knob, which develops on the apex of a hyphal branch. RNA was isolated from knobs and mycelium and hybridized to a cDNA array containing probes of 2822 EST clones of *M. haptotylum*. Despite the fact that the knobs and mycelium were grown in the same medium, there were substantial differences in the patterns of genes expressed in the two cell types. In total, 23·3 % (657 of 2822) of the putative genes were differentially expressed in knobs versus mycelium. Several of these genes displayed sequence similarities to genes known to be involved in regulating morphogenesis and cell polarity in fungi. Among them were several putative homologues for small GTPases, such as *rho1*, *rac1* and *ras1*, and a rho GDP dissociation inhibitor (*rdi1*). Several homologues to genes involved in stress response, protein synthesis and protein degradation, transcription, and carbon metabolism were also differentially expressed. In the last category, a glycogen phosphorylase (*gph1*) gene homologue, one of the most upregulated genes in the knobs as compared to mycelium, was characterized. A number of the genes that were differentially expressed in trap cells are also known to be regulated during the development of infection structures in plant-pathogenic fungi. Among them, a *gas1* (*mas3*) gene homologue (designated *gks1*), which is specifically expressed in appressoria of the rice blast fungus, was characterized.

INTRODUCTION

The nematophagous fungi comprise more than 200 species of taxonomically diverse fungi that all share the ability to infect and kill living nematodes (Barron, 1977). The interest in studying these fungi arises from their potential use as biological control agents for nematodes that are parasitic on plants and animals (Larsen, 2000). Other reasons for the continued fascination with nematophagous fungi are their remarkable morphological adaptations and their dramatic infection of nematodes. The nematode-trapping fungi are particularly interesting in this respect. These fungi enter the parasitic stage by developing specific morphological structures called traps. The traps develop from hyphal branches and they can be formed either spontaneously or be induced in response to signals from the environment, including peptides and other compounds secreted by the host nematode (Dijksterhuis *et al.*, 1994). There is a large variation in the morphology of trapping structures, even between closely related species (Ahrén *et al.*, 1998; Barron, 1977). In some species, the trap consists of an erect branch that is covered with an adhesive material. In other species, such as in the well-studied *Arthrobotrys oligospora*, the trap is a complex three-dimensional net. A third type of trap is the adhesive knob. The knob is a morphologically distinct cell, often produced on the apex of a slender hyphal stalk. A layer of adhesive polymers covers the knob, which is not present on the support stalks. Finally, there are some species of nematode-trapping fungi that capture nematodes in mechanical traps called constricting rings (Barron, 1977).

†Present address: European Bioinformatic Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK.

‡These authors contributed equally to this work.

Abbreviations: aRNA, antisense RNA; EST, expressed sequence tag.

Microarray raw data are available at the EBI-EMBL ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) (accession no. E-MEXP-250).

A comparison of the primary structure of the *M. haptotylum* gph1 protein with the phosphorylases of yeast (*Saccharomyces cerevisiae*) and rabbit (*Oryctolagus cuniculus*) is available in Supplementary Fig. S1a, and a phylogeny of glycogen phosphorylases in Supplementary Fig. S1b, with the online version of this paper at http://mic.sgmjournals.org/.

The ultrastructure of the nematode-trapping structures has been examined extensively. These studies have shown that despite the large variation in morphology, the adhesive types of trap (branches, nets and knobs) have a unique ultrastructure that clearly distinguishes them from vegetative hyphae (Dijksterhuis *et al.*, 1994). One feature, which is common to all of these traps, is the presence of numerous cytosolic organelles, the so-called dense bodies. Although the function of these organelles is not yet clear, the fact that they exhibit catalase and D-amino acid oxidase activity indicates that the dense bodies are peroxisomal in nature (Dijksterhuis *et al.*, 1994). Another feature common to the trap cells is the presence of extensive layers of extracellular polymers, which are thought to be important for attachment of the traps to the surface of the nematode (Tunlid *et al.*, 1991).

In this study we have analysed the global pattern of gene expression in knobs and mycelium of the fungus *Monacrosporium haptotylum* (syn. *Dactylaria candida*) (Fig. 1). The advantage of using *M. haptotylum* is that during growth in liquid cultures with heavy aeration, the connections between the traps (knobs) and mycelium can be broken easily and the knobs can be separated from the mycelium by filtration (Friman, 1993). The isolated knobs retain their function as infection structures: they can 'capture' and infect nematodes. In order to characterize overall gene expression, we have conducted expressed sequence tag (EST) analysis with 8466 clones from four different cDNA libraries. These include libraries made from mycelium and knobs, and two libraries from knobs infecting the nematode *Caenorhabditis elegans*. Since the genome of *C. elegans* has been sequenced, transcripts expressed by the fungus were easily separated from those of the host. From the EST information assembled, we obtained a unique set of 3036 clones, of either fungal or worm origin, which was used for the construction of cDNA microarrays. Here we describe the analysis of gene expression patterns in knobs and mycelium.

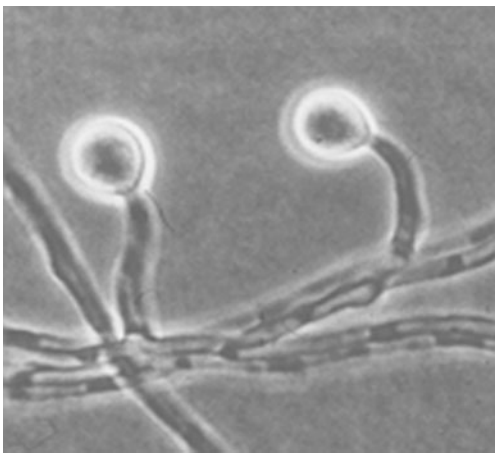

Fig. 1. Trapping structures of *Monacrosporium haptotylum*. The knob cell is a spherical cell formed on the apex of a hyphal branch. The picture is reproduced from Nordbring-Hertz *et al.* (1995), courtesy of Birgit Nordbring-Hertz and IWF, Göttingen.

METHODS

Culture of organisms and infection experiments. *Monacrosporium haptotylum* (CBS 200.50) was grown in aerated 5 l liquid cultures containing 0·01 % soya peptone, 0·005 % phenylalanine and 0·005 % (w/v) valine. The knobs were separated from the vegetative mycelium by filtering the culture through a nylon mesh (10 μm), and collected from the filtrate on a 1·2 μm pore-size membrane filter (Friman, 1993). The vegetative mycelium and knob fractions were immediately frozen in liquid nitrogen and stored at −80 °C. *C. elegans* (strain N2) was grown using standard protocols (Brenner, 1974). To partly synchronize the nematodes, they were treated with 0·25 % KOH and 2·5 % NaOCl (Emmons *et al.*, 1979). L1 larvae were incubated on water agar plates with freshly isolated knobs from *M. haptotylum*, and the infection was followed under a light microscope. After both 4 h and 24 h, the infected nematodes were collected by flooding the plates with water. The samples were centrifuged and frozen in liquid nitrogen.

RNA isolation and construction of cDNA libraries. Total RNA was extracted from the vegetative mycelium, the knobs and infected *C. elegans* using the RNeasy plant mini kit (Qiagen) with the guanidinium isothiocyanate buffer provided. The integrity of the total RNA was checked by denaturating formaldehyde agarose gel electrophoresis and the quantity was assessed by UV spectrophotometry. Four unidirectional cDNA libraries were constructed from total RNA using the PCR-based SMART cDNA Library Construction kit (BD Clontech). After ligation of size-fractionated cDNA (>100 bp) into the λTriplEx2 arms (BD Clontech) and lambda phage packaging (GigaPack III Gold extract; Stratagene), the primary titres for the unamplified libraries were determined to be $5{\cdot}9\times10^6$ p.f.u. for mycelium, $1{\cdot}6\times10^6$ p.f.u. for knob, $2{\cdot}5\times10^6$ p.f.u. for the 'early' stage of *C. elegans* infection (4 h) and $2{\cdot}9\times10^6$ p.f.u. for the 'late' stage of *C. elegans* infection (24 h). In order to amplify the libraries, phage suspensions were used to transduce *Escherichia coli* XL-1 Blue, which resulted in titres of more than $5{\cdot}0\times10^9$ p.f.u. ml^{-1}. The λTriplEx2 libraries were converted to plasmid libraries by *in vivo* excision in *E. coli* BM25.8 (Johansson *et al.*, 2004). PCR amplification followed by gel electrophoresis demonstrated that more than 90 % of the randomly collected clones contained insert. Based on the analysis of 96 individual clones from each of the four libraries, the average sizes of the inserts in the mycelium, knob, 4 h infection and 24 h infection cDNA libraries were found to be approximately 1100, 1000, 750 and 1200 bp, respectively. The size of the transcripts ranged from 120 to 2400 bp in all four libraries.

EST analysis. Plasmid clones were randomly collected from each of the four libraries. The colonies were transferred to 96-well plates, and the bacterial lysates were used as starting material for PCR amplification (Johansson *et al.*, 2004). PCR was performed using pTriplEx2-specific universal primers P104 (5′-GGGAAGCGCGCC-ATTGTGTT-3′) and P105 (5′-AGTGAGCTCGAATTGCGGCC-3′) in a total volume of 10 μl, followed by assessment of the size and quality of the products by gel electrophoresis. Partial nucleotide sequences of amplified cDNA inserts were determined using the P104 primer and the dideoxy chain-termination method employing the BigDye Terminator Cycle Sequencing kit (Applied Biosystems). The sequencing reaction products were analysed using either a 377 or 3100 ABI sequencer (Applied Biosystems). cDNA clones representing genes that are discussed in more detail were verified and completely sequenced by using universal and gene-specific primers (Table 1).

Table 1. Gene-specific primers used in this study

Gene	EST	Description	Primer ID	Sequence
rho1	CN800249	GTPase Rho1	P337	5′-GAAGTTGATGGGGAGCATGT*†
			P338	5′-TGCAAGACCTCAGAAATCCA*†
			S219	5′-GCCGCCAAGTACCTAGAGTG*
			S220	5′-CCGTGCTGCCCTTACTACTT*
rac1	CN795728	GTPase Rho3	P351	5′-CCTCCATTTCAGCTTCCAAA*
			P352	5′-TTCCGTCGACCATGACACTA*
			S229	5′-AGTGGTTCCCGGAAATTAGC*
			S230	5′-GCGCCAAATGTTCCAATTAT*
			S231	5′-GCGACACCATTTGTGAACTG*
ras1	CN799032	Ras homologue	P333	5′-GGAGAGGGTGGTGTTGGTAA*†
			P334	5′-TTGCCTCGTGTCCTGTGATA*†
			S214	5′-CGACCTACTGGGAGAACGTG*
			S215	5′-AACCTCAGCAAAGGATAGGG*
gks1	CN800713	GAS1 (MAS3) protein	P341	5′-AGACAGACACCACCCGATTC*†
			P342	5′-TTGACCTGGTGAAGGGTCAT*†
			S224	5′-GCGGTGGTAACCAGAACACT*
			S225	5′-AAAGTTGACGACGGAAGGAA*
gph1	CN795977	Glycogen phosphorylase	P331	5′-GGACATCCTGAGCCAAACAT*
			P332	5′-GCAAGCTCATCGCAATTGTA*
			F04	5′-AGTCCGCTTGTGTCTTGGAC*
			F05	5′-CCCCGACACGAGATTACTGT*
			D06	5′-TGGTCTGTTCCACTTGTCCA*
			R04	5′-GCCATCTCCGAGGAGTAACA*
			R05	5′-CCACCGTTGAGAACGAACTT*
			D08	5′-CTCGGCGGGTTTACTTTCTT*
			F10	5′-CCCAGGTGTCTTCAGTGGTT*
			L08	5′-GCTGTGTCTGAGGTCGTCAA*†
			R05	5′-CCACCGTTGAGAACGAACTT*†
gpd1	CN795977	Glyceraldehyde-3-phosphate dehydrogenase	H1L	5′-GCCACCAACATTATCCCATC†
			H1R	5′-GTTGGGACACGGAAAGACAT†
	CN801309	Cuticle-degrading serine protease	P307	5′-CGCTGTTAACACCCACACAG*†
			P308	5′-CACGGTAGCCTTGTCGAAAT*†
	CN800950	Alkaline serine protease	P315	5′-GACACGGAACACACACTGCT*†
			P316	5′-TTTTGATGCTCGCTGATGAC*†
	CN800618	β-1,3 exoglucanase	P311	5′-CAATATTTGCCGCTTCATCA*
			P312	5′-TGCACTATTGGGTGAACCAG*
	CN800820	Trichothecene 3-*O*-acetyl-transferase	P313	5′-TTCAGTAACCGAGCCCATTC*†
			P314	5′-CGCGATTGTTTGTAACATGG*†
	CN794863	Immunoreactive spherule cell-wall protein	P329	5′-CTTGACCAGATCCCATCCTG*†
			P330	5′-CGGTGTTCTCGTACTCAGCA*†

*Primers used for sequencing.

†Primers used for real-time RT-PCR.

Base-calling of DNA sequencer traces was conducted using the PHRED program with the quality level set to 20 (Ewing *et al.*, 1998). Subsequently, poly(A) and vector sequence information was removed and the sequences from all three stages were assembled into contigs using the Contig Assembly Program (CAP) (Huang, 1992). After translation into all three forward frames, similarity analyses were performed by FASTA searches (Pearson, 1994) against the GenBank non-redundant (nr) protein database. Based on sequence homology, the ESTs were assigned functional roles and EC numbers (where applicable). Functional roles were classified into categories according to catalogues used for the *Saccharomyces cerevisiae* genome provided by the Munich Information Centre for Protein Sequences (MIPS) (Mewes *et al.*, 2000). All EST information was stored in an SQL database (MySQL) on a Linux (Red Hat 6.2) platform, and the data were processed using the PHOREST tool (Ahren *et al.*, 2004). Of 8466 EST clones analysed, 7115 high-quality ESTs (>99 bp) were deposited in the GenBank dbEST database and are available by accession numbers CN794415 to CN801529.

To separate fungal sequences from worm sequences, a local BLASTN search was performed against 22 055 *C. elegans* cDNA sequences downloaded from the ENSEMBLE website (http://www.ensembl.org/

Caenorhabditis_elegans/). Sequences with an E value $<1{\cdot}0\times10^{-4}$ were annotated as *C. elegans* and sequences with an E value $>1{\cdot}0\times10^{-4}$ were annotated as fungal. Sequences of <50 bp were annotated as unknown.

Construction of cDNA arrays. A single EST cDNA clone was selected to represent each assembled sequence (i.e. putatively unique transcript). For an assembled contig represented by multiple ESTs, the rule followed was to select the clone with the most extensive DNA sequencing read length. These clones were collected from the master EST library stock contained as bacterial lysate plates and aliquots were transferred into new 96-well plates. Plasmid inserts were amplified by PCR, purified and concentrated as previously described (Johansson *et al.*, 2004). In total, 3518 clones representing approximately 3036 contigs were processed successfully. cDNA PCR products were printed on CMT-GAPSII-coated slides (Corning Glass) using a 16-pin configured MicroGrid II array printer (BioRobotics) controlled by the MicroGrid TAS Application Suite (version 2.2.0.6). The general design of the microarrays was two identical tool arrays, each containing 16 blocks (sub-arrays) and with each clone replicated twice within each block, which provided on-chip quadruplicates for each clone. Additionally, various positive and negative control reporters were replicated in various positions within each block, such as positive homologous controls, that is, *M. haptotylum* and *C. elegans* genes expected to be highly expressed. In addition, eight different heterologous and commercial PCR reporters (ArrayControl; Ambion) for which complementary RNAs were spiked into the amplification and labelling process of targets were also included in the blocks. Each control reporter was present in up to 64 per-chip replicates. DNA was cross-linked to the slides by baking at 80 °C for 2 h in glass jars, followed by UV cross-linking at 90 mJ cm^{-2}. Spot and print quality were assessed by staining with POPO3 dye (Molecular Probes), and subsequent scanning showed that no spots were missing or joined on the slides.

Microarray analysis. To compare the expression pattern of genes in mycelium and knobs of *M. haptotylum*, our microarray experiments were designed as two-samples comparisons (Kerr & Churchill, 2001; Wu, 2001) (i.e. knobs versus mycelium) using three independent biological replicates, which also included technical and dye-swapped control hybridizations. Nine cDNA microarrays were hybridized; four of these were hybridizations of knobs versus mycelium and five were mycelium against mycelium. cDNA probes were prepared from total RNA isolated from vegetative mycelium and knobs, grown under identical conditions to those used in the construction of the cDNA libraries. Total RNA was extracted according to Stiekema *et al.* (1988) and purified further on an RNeasy plant mini kit spin column (Qiagen). To obtain the required amounts of RNA from limited amounts of starting material, two rounds of antisense RNA (aRNA) amplification were performed using the MessageAmp kit (Ambion). The integrity of aRNA was analysed by denaturating formaldehyde agarose gel electrophoresis and quantified using the RiboGreen RNA Quantification kit (Molecular Probes). The aRNA samples were divided into aliquots and stored at −80 °C. A total of 2 μg aRNA for each target, including ArrayControl RNA exogenous spikes (Ambion), was labelled (Cy3 or Cy5) using the CyScribe post-labelling kit (Amersham Biosciences). The labelled cDNA was purified with the CyScribe GFX Purification kit (Amersham Biosciences). The slides were hybridized, rinsed and scanned as previously described (Johansson *et al.*, 2004). The microarray raw data are available at the EBI-EMBL ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) (accession no. E-MEXP-250).

Statistical analysis. After scanning, data images were inspected manually and low-quality spots were excluded before further analysis. Reporters of unknown origin and related to *C. elegans* were excluded, and only data for 2822 reporters representing fungal genes were extracted. For those spots remaining, the raw fluorescence intensities for each channel on each slide were collected. The mean background fluorescence was calculated for each channel. After local background correction for each spot, the reporters yielding intensities below twice the mean background were excluded and the fluorescence for the remaining reporters was multiplied to give a common channel mean of 5000 fluorescence units. As a result, data for 2534 fungal genes remained in the dataset. The statistical approach we used, the mixed-model analysis of variance (ANOVA), served two purposes: (1) normalization of the data to remove systemic biases that may have affected all genes simultaneously, such as differences in the amount of RNA that was labelled for a particular replicate of a treatment; (2) assessment of the contribution of biological and experimental sources of error to the variation in expression of each individual gene (Wolfinger *et al.*, 2001). This procedure uses differences in normalized expression levels, rather than ratios, as the unit of analysis of expression differences. We subjected the corrected $\log_2$-transformed measures (y_{gij}) for the fungal gene g ($g=1,\ldots,2534$) which included scores for 107 445 fungal spot measures to a normalization model of the form $y_{gij}=\mu+A_i+D_j+(A\times D)_{ij}+\varepsilon_{gij}$, where μ is the sample mean, A_i is the effect of the ith array ($i=1$–9), D_j is the effect of the jth dye (Cy3 or Cy5), $(A\times D)_{ij}$ is the array–dye interaction (channel effect), and ε_{gij} is the stochastic error. We then subjected the residuals from this model, which can be regarded as crude indicators of relative expression level (and which are referred to in the text as 'normalized gene expression levels') to 2534 gene-specific models of the form $r_{ijlm}=\mu+A_i+T_l+D_j+\varepsilon_{ijlm}$, where T_l is the lth tissue (knob or mycelium; 1 degree of freedom, d.f.). In the gene models, which were fitted using PROC MIXED in SAS (SAS/STAT software version 8, SAS Institute), the A_i variable controls for spot effects and is random (8 d.f., leaving 7 d.f. for the residual error).

Real-time quantitative PCR. aRNA was reverse transcribed into cDNA using the SYBR Green RT-PCR Reagents kit (Applied Biosystems). The reverse transcription (RT) reactions were normalized to contain equivalent amounts (2 μg) of aRNA. RT reactions were primed by random hexamers and were carried out in a total volume of 10 μl according to the manufacturer's instructions. The cDNA product was treated with 2 units of RNase H (Invitrogen) for 20 min at 37 °C, followed by heat inactivation at 75 °C for 10 min.

Gene-specific oligonucleotide primers were designed by using the web-based primer picking service (Primer3) (http://fokker.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) (Rozen & Skaletsky, 2000). The following criteria were applied: melting temperatures 58–60 °C, primer lengths 20–22 nt, GC content 40–80 %, PCR amplicon length 100–200 bp (Table 1). The PCR reactions were performed with an Mx3000P Real-Time PCR system (Stratagene) using SYBR Green PCR Master Mix (Applied Biosystems) to monitor dsDNA synthesis. The reactions contained 12·5 μl 2× SYBR Green PCR Master Mix (including AmpliTaq Gold DNA Polymerase, dNTPs with dUTP, SYBR Green I Dye, Passive Reference and optimized buffer components), 1 μl cDNA (diluted 1 : 10), 1 μl gene-specific forward and reverse primers (0·5 μmol l^{-1} of each), in a final volume of 25 μl. The following thermal profile was used for all PCRs: an initial denaturation step at 95 °C for 10 min, followed by 100 cycles at 95 °C for 30 s, 58 °C for 1 min, and 72 °C for 1 min, and finally extended elongation at 72 °C for 30 s. Reactions were set up in triplicate, including a control with no template. Amplification of gene-specific products was analysed by melting-curve analysis followed by agarose gel electrophoresis (cf. Fig. 3).

The real-time RT-PCR data were analysed using relative quantification according to the $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen, 2001) (Stratagene software for the Mx3000P Real-Time PCR instrument). Briefly, the relative quantity is presented as the fold change in gene expression normalized to an internal reference (control) gene in knobs relative to the normalized expression levels in the mycelium. The

purpose of the internal control gene is to establish that the RNA targets are reverse transcribed and subsequently amplified with similar efficiency in each reaction. The putative house-keeping gene glyceraldehyde-3-phosphate dehydrogenase (*gpd1*) served as reference. The amplification efficiency (Eff) of each target gene was compared to that of *gpd1* at different template concentrations, and was found to be similar (81·6 %–95·1 %) to that of *gpd1* (not shown).

Phylogenetic analysis. To construct the phylogenetic trees of GTPases, gEgh16 homologues and glycogen phosphorylases, homologous sequences were retrieved from the GenBank nr protein database (Benson *et al.*, 2004) using BLAST searches (Altschul *et al.*, 1990). Protein alignments were made using CLUSTAL W (Thompson *et al.*, 1994). Phylogeny was done using the Neighbour-Joining method (Saitou & Nei, 1987) with 500 bootstrap replications implemented in the MEGA 2.1 program (Kumar *et al.*, 2001).

RESULTS

EST analysis

Four directional cDNA libraries were constructed from mycelium, knobs, and knobs infecting *C. elegans* for 4 and 24 h, respectively. The knobs were isolated from liquid-grown mycelium of *M. haptotylum* and then incubated with *C. elegans* on water agar plates. After 4 h of infection, a number of the knobs had adhered and had started to penetrate the cuticle of *C. elegans*. Approximately 30–35 % of the nematodes were non-motile and were considered to have been killed by the fungus. After 24 h of infection, approximately 80–85 % of the nematodes were immobile, and fungal hyphae were growing inside the infected nematodes.

In total, 8466 ESTs were obtained from the four cDNA libraries. The sequences were assembled into 3121 contigs that putatively represent unique genes/transcripts (Table 2). Of the assembled sequences, 723 were represented in the mycelium library, 574 in the knob library, 682 in the 4 h infection library and 1531 in the 24 h infection library. Between 56 and 71 % of the contigs were singletons: they were represented by only one EST clone. Between 5 and 37 % of the assembled sequences displayed a high degree of similarity (FASTA score >299) to sequences in the GenBank nr protein database (Table 2). The fraction of ESTs showing high-score homology was largest in the 4 h infection library. This is due to the fact that this library contained a large proportion of *C. elegans* sequences: 67 % of the contigs in this library were identified as being of *C. elegans* origin, while 21 % were of fungal origin and 12 % of unknown origin. The fraction of *C. elegans* transcripts in the 24 h library was considerably lower (3·7 %), whereas 88 % were of fungal and 8·2 % of unknown origin. A large fraction (38–60 %) of the assembled sequences showed no homology (orphans) to protein sequences in the GenBank nr protein database, and the proportion of orphans was highest in the knob (58 %) and 24 h infection (60 %) libraries.

Expression of genes in knobs and mycelium

To allow more detailed studies on the expression of genes in *M. haptotylum* and during its infection of *C. elegans*, we constructed a cDNA microarray from the set of EST clones analysed. In total, 3518 clones were amplified and spotted on the array, 2822 of fungal, 540 of *C. elegans* and 156 of unknown origin. In the hybridization experiments, we found 2534 fungal reporters (90 % of the total number of fungal reporters present on the array) to produce signals above the two-fold average background intensity on one or more slides. To determine which of these genes differed significantly in expression between knobs and mycelium, the microarray data were analysed using the mixed-model ANOVA (Wolfinger *et al.*, 2001). Direct visualization of the significance and magnitude of differences in gene expression between the knobs and mycelium was possible by the use of a volcano plot (Fig. 2). This plot contrasts significance, $-\log_{10}(P)$, on the y axis against fold change in expression between knob and mycelium on the x axis. The distribution along the x axis shows the expression difference, whereas the distribution along the y axis shows the degree of significance, with 2, 3, etc. representing P values of 10^{-2}, 10^{-3}, etc. A significant observation made from this plot is that apparently more genes are down-regulated than upregulated in the knobs relative to mycelium. Choosing a significance level of $P<0{\cdot}001$, the number of down-regulated genes was 299 (10·6 % of the total number of fungal clones) and the number of upregulated genes was 193 (6·8 %) (Table 3).

To validate the microarray data, the identities of 16 reporters were confirmed by resequencing (Table 4). In addition, the expression levels of eight of these genes were analysed by quantitative real-time RT-PCR (Fig. 3, Table 5). Although there were some differences in the estimated fold values, the overall patterns of expression were similar using the DNA microarray and real-time RT-PCR.

When knobs and mycelium were compared, there was a large difference in the functional distribution of genes being differentially regulated (Fig. 4). A significant proportion of the genes that had putative roles in 'transcription', 'cellular transport and transport mechanisms', 'cellular communication or signal transduction', 'cell rescue, defence, cell death and ageing' and 'cell growth, cell division and DNA synthesis' were expressed at lower levels in the knobs than in the mycelium. Some of the genes that were differentially expressed in the two cell types are listed in Table 4, and cDNA clones representing a few of these were selected for further characterization by full-length sequencing.

DISCUSSION

Morphogenesis and polarity

The trap in *M. haptotylum* is a spherical cell, which develops at the tip of an apical growing hyphal branch (Fig. 1). This change in morphology represents a shift in the polarity of the cells. In fungi, the actin cytoskeleton provides the structural basis for establishing and modifying cell polarity

Table 2. Number of EST clones analysed in each cDNA library

The distribution of EST clones was according to various ranges of sequence homology scores after sequence comparison with the GenBank (nr) protein database. Assembled sequences, putatively representing different transcripts, refer to contigs (assembled continuous sequences) and singletons after contig assembly (Huang, 1992). The Singleton percentages represent the fraction of singletons compared to the total number of contigs, and indicate the level of redundancy.

Library	Total ESTs (>1 bp)	Assembled sequences		Sequence homology (FASTA/GenBank) scores (%)			
		Contigs	Singletons	Orphans	Low (<100)	Moderate (100–299)	High (>299)
Mycelium	1937	723	508 (56%)	45	15	28	12
Knob	1809	574	368 (64%)	58	11	19	12
Infection (4 h)	1522	682	483 (71%)	38	4	21	37
Infection (24 h)	3198	1531	1076 (70%)	60	14	22	5
All libraries	8466	3121	2008 (64%)				

(Pruyne & Bretscher, 2000a, b). During apical growth, the actin cytoskeleton polarizes growth by assembling into cortical patches and actin cables at the tip. During isotropic growth, the proteins of the cap are more diffusely distributed and actin cables form a meshwork. Several genes are differentially expressed in knobs and mycelium, and these display significant sequence similarities to genes known to be involved in modifying cell polarity in fungi. Among them were two down-regulated homologues to the actin-binding proteins profilin (EST clone CN795458, cDNA AY635995, average fold change on microarray −1·26 ($\log_2$, knobs versus mycelium)) and cofilin (EST clone CN795276, cDNA AY635996, fold change on microarray −1·87). Profilin is a small actin-binding protein, which was originally identified as an actin monomer sequestering protein that can inhibit the growth of actin filaments and prevent polymerization. Studies have also indicated that profilin can have a role in promoting actin polymerization (Ayscough, 1998). Cofilin binds to actin monomers and its activity is important for rapid turnover of actin filaments in *S. cerevisiae* (Lappalainen & Drubin, 1997).

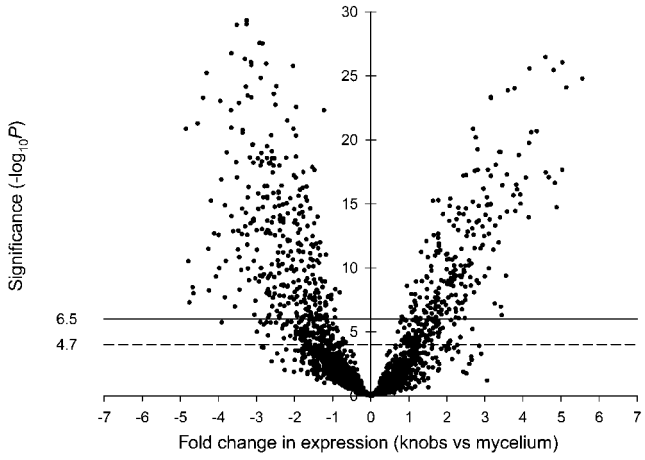

Fig. 2. Volcano plot of significance against fold change in gene expression, in knobs versus mycelium. Each point represents a single gene analysed by the mixed model ANOVA (Wolfinger *et al.*, 2001). The negative $\log_{10}$ of the *P* value is plotted against the difference between least-square means of $\log_2$-normalized expression values. Highly significant values are located towards the top and small expression differences are located at the centre of each plot. The horizontal dashed line represents the test-wise threshold of $P=0{\cdot}05$ (corresponding to a Bonferroni corrected value of $-\log_{10}=4{\cdot}7$), and the solid line represents a threshold of $P=0{\cdot}001$ ($-\log_{10}=6{\cdot}5$).

The rho and ras family of small GTPases, including rho, rac and cdc42, play an important role in regulating the actin part of the cytoskeleton (Pruyne & Bretscher, 2000a), and these molecules act as molecular switches through their ability to hydrolyse GTP. In the GTP-bound form, these signalling proteins are active and exert a positive signal on proteins associated with polarized growth. We identified and characterized three small GTPases, including a rho1, a rac and a ras homologue (Fig. 5) that were differentially expressed in knobs and mycelium. Phylogenetic analysis showed that the EST clone CN800249 (fold change on

Table 3. Number of genes showing significant differential regulation when knobs and mycelium of *M. haptotylum* were compared

The number of genes with a *P* value of less than the indicated level, as determined by mixed model ANOVA (Wolfinger *et al.*, 2001), is shown. $P<0{\cdot}001$ corresponds to a Bonferroni value of $3{\cdot}5\times10^{-7}$, $P<0{\cdot}01$ to $3{\cdot}5\times10^{-6}$ and $P<0{\cdot}05$ to $1{\cdot}8\times10^{-5}$. The percentage fraction of fungal genes regulated is given in parentheses (2822 reporters in total).

	Significance level (*P*)		
	0·001	0·01	0·05
Upregulated	193 (6·8)	234 (8·3)	259 (9·2)
Down-regulated	299 (10·6)	348 (12·3)	398 (14·1)
Total	492 (17·4)	582 (20·6)	657 (23·3)

Table 4. Selected genes differentially expressed in knobs versus mycelium of *M. haptotylum*

Clones in bold have been verified by resequencing. Fold change represents the mean fold change (log_2) values of expression levels in knobs (trap cells) compared to mycelium; standard error is given in parentheses. All genes are differentially regulated at a significance level of $P<0{\cdot}001$ (Bonferroni $3{\cdot}5\times10^{-7}$; cf. Table 3).

EST clone	GenBank GI	GenBank homologue description	Score	Fold change	Function
Cell growth, cell division and DNA synthesis					
CN795465	113277	Actin gamma (*Emericella nidulans*)	633	−1·23 (0·12)	
CN795458	19114739	Profilin (*Sc. pombe*)	230	−1·26 (0·18)	Actin-associated protein
CN795276	11596089	Cofilin (*Zygosaccharomyces rouxii*)	235	−1·87 (0·26)	Actin-associated protein
CN794706	7673696	*β*-Tubulin (*Leptosphaeria maculans*)	409	−1·43 (0·25)	
CN795822	729011	Calmodulin (*Neurospora crassa*)	419	−1·22 (0·23)	
CN795399	461631	BMH2 protein (*S. cerevisiae*)	414	−1·41 (0·17)	14-3-3-like protein
CN797412	8218222	Multifunctional cyclin-dependent kinase PHO85 (*N. crassa*)	501	−2·64 (0·26)	Cyclin dependent protein kinase
CN798834	9296975	Glucosamine-phosphate *N*-acetyltransferase GNA1 (*Candida albicans*)	251	−1·27 (0·17)	UDP-*N*-acetylglucosamine biosynthesis
CN794706	6321670	Prohibitin PHB2 (*S. cerevisiae*)	162	2·06 (0·19)	Mitochondrial protein regulate replicative lifespan
Cell rescue, defence, cell death and ageing					
CN794866	11139253	Mitochondrial thioredoxin peroxidase PRX1 (*S. cerevisiae*)	559	−1·42 (0·21)	
CN795223	11135375	Thioredoxin (*Coprinus comatus*)	368	−2·41 (0·40)	
CN795856	16416067	Related to hsp70 protein (*N. crassa*)	248	−3·57 (0·56)	HSP70 family
CN796871	171388	DDR48 stress protein (*S. cerevisiae*)	211	−2·71 (0·27)	Flocculent specific protein
CN800033	16943775	Putative cyclophilin (*Pleurotus ostreatus*)	506	−3·16 (0·13)	
CN800928	3869176	Cyclophilin (*Arthroderma benhamiae*)	250	1·87 (0·28)	
CN799935	17228073	Peptidyl-prolyl cis-trans isomerase (*Nostoc* sp. PCC 7120)	346	1·29 (0·20)	
CN799620	6323840	Heat-shock protein HSC82 (*S. cerevisiae*)	450	1·80 (0·26)	HSP90 family
Cellular communication/signal transduction					
CN800249	13432032	GTPase Rho1 (*A. fumigatus*)	533	−2·03 (0·14)	Rho GTPases
CN795728	13432036	GTPase Rho3 (*A. fumigatus*)	341	−1·32 (0·23)	Rho GTPases
CN794767	6320066	Rho GDP dissociation inhibitor Rdi1p (*S. cerevisiae*)	477	−2·01 (0·18)	GDP dissociation inhibitor
CN799032	3929359	Ras homologue (*Col. trifolii*)	305	2·67 (0·23)	Ras GTPases
Cellular organization					
CN796915	2498762	Peroxisomal membrane protein PMP30B (*C. boidinii*)	158	0·72 (0·15)	Peroxisome biogenesis and organization
Cellular transport and transport mechanisms					
CN795839	546371	GDP dissociation inhibitor Gdi1 (*S. cerevisiae*)	187	−2·64 (0·20)	RAB GDP dissociation inhibitor
CN794610	5881950	Putative mitochondrial phosphate carrier protein (*Tuber magnatum*)	317	−1·20 (0·16)	Mitochondrial Carrier (MC) family
CN800096	14700035	Nuclear transport factor 2 (*E. nidulans*)	222	−2·02 (0·25)	
CN800122	728908	Plasma membrane ATPase (*Ajellomyces capsulatus*)	472	−3·14 (0·19)	Membrane ATPase
CN800520	16944648	Related to ATP-binding cassette transporter protein YOR1 (*N. crassa*)	348	1·60 (0·28)	ABC transporter
Metabolism					
CN795758	1084868	Subunit of trehalose-6-phosphate synthase/phosphatase complex; TPS3 (*S. cerevisiae*) (EC 2.4.1.15)	282	−3·40 (0·40)	Trehalose biosynthesis
CN794710	2114495	UDP-glucose dehydrogenase (*Drosophila melanogaster*) (EC 1.1.1.22)	305	−2·60 (0·23)	Carbohydrate metabolism
CN800386	7493924	Alpha-mannosidase *E. nidulans* (EC 3.2.1.24)	422	−1·67 (0·22)	Carbohydrate metabolism
CN795087	15278113	Cystathionine gamma-lyase (*N. crassa*) (EC 4.4.1.1)	244	−1·58 (0·24)	Amino acid biosynthesis
CN795017	15072557	Homoaconitase (*A. fumigatus*) (EC 4.2.1.36)	141	−2·52 (0·24)	Lysine biosynthesis
CN796046	129302	D-amino acid oxidase (*Nectria haematococca*) (EC 1.4.3.3)	289	−3·09 (0·29)	Amino acid metabolism
CN795379	1078614	Stearoyl-CoA desaturase (*Aj. capsulata*) (EC 1.14.99.5)	352	−3·38 (0·29)	Fatty acid desaturation
CN795279	1168273	Acyl-CoA-binding protein (ACBP) (*D. melanogaster*)	298	−1·93 (0·19)	Lipid metabolism
CN795977	130174	Glycogen phosphorylase GPH1 (*S. cerevisiae*) (EC 2.4.1.1)	338	4·36 (0·23)	Glycogen degradation
CN800618	11359357	Beta-1,3 exoglucanase (*Trichoderma harzianum*) (EC 3.2.1.−)	187	1·41 (0·25)	Glucan metabolism
CN800820	15054404	Trichothecene 3-*O*-acetyltransferase (*Fusarium sporotrichioides*)	154	2·00 (0·23)	Secondary metabolism
CN794616	8134725	Spermidine synthase (*N. crassa*) (EC 2.5.1.16)	643	1·46 (0·20)	Spermidine and pantothenate biosynthesis

Table 4. cont.

EST clone	GenBank GI	GenBank homologue description	Score	Fold change	Function
CN795989	19115006	Putative nicotinate phosphoribosyltransferase (*Sc. pombe*) (EC 2.4.2.11)	345	1·08 (0·18)	Nicotinate and nicotinamide metabolism
CN797536	6324126	Peroxisomal 2,4-dienoyl-CoA reductase SPS19 (*S. cerevisiae*) (EC 1.3.1.34)	195	1·05 (0·17)	Fatty acid degradation
CN796073	41629676	Fatty acid transporter and very long-chain fatty acyl-CoA synthetase FAT1 (*S. cerevisiae*) (EC 6.2.1.3)	257	−2·30 (0·21)	Fatty acid biosynthesis
Energy					
CN800416	14699983	Triose phosphate isomerase (*Paracoccidioides brasiliensis*) (EC 5.3.1.1)	269	−1·81 (0·37)	Glycolysis and gluconeogenesis
CN796696	2494634	Glyceraldehyde-3-phosphate dehydrogenase (*Aspergillus niger*) (EC 1.2.1.12)	459	−1·59 (0·24)	Glycolysis and gluconeogenesis
CN797534	13925873	Enolase (*A. fumigatus*) (EC 4.2.1.11)	574	−2·07 (0·25)	Glycolysis and gluconeogenesis
CN797352	9955865	Fructose-1,6-bisphosphate aldolase (*Aspergillus oryzae*) (EC 4.1.2.13)	390	−2·53 (0·33)	Glycolysis and gluconeogenesis
CN797016	11257242	Probable transaldolase (*Sc. pombe*) (EC 2.2.1.2)	265	−2·17 (0·23)	Pentose-phosphate pathway
CN795968	4029338	Malate dehydrogenase (*Piromyces* sp. *E2*) (EC 1.1.1.37)	428	−2·53 (0·19)	Tricarboxylic acid pathway
CN794567	9695396	NADH dehydrogenase subunit 1 (*Phytophthora infestans*) (EC 1.6.5.3)	152	−4·79 (0·52)	Oxidative phosphorylation
CN796765	548389	NADH dehydrogenase chain 4L (*T. rubrum*) (EC 1.6.5.3)	247	1·77 (0·25)	Oxidative phosphorylation
CN801000	1334599	Ubiquinol : cytochrome *c* oxireductase (*Ma. grisea*) (EC 1.10.2.2)	616	3·36 (0·31)	Oxidative phosphorylation
CN794469	1351364	Ubiquinol : cytochrome *c* reductase complex (EC 1.10.2.2)	255	−1·89 (0·35)	Oxidative phosphorylation
CN795459	117840	Cytochrome b (*E. nidulans*) (EC 1.10.2.2)	484	2·78 (0·24)	Oxidative phosphorylation
CN799786	14269417	Cytochrome-c oxydase chain VIIc-like protein (*O. novo-ulmi*) (EC 1.9.3.1)	156	−1·68 (0·14)	Oxidative phosphorylation
CN801258	584806	ATP synthase alpha chain (*N. crassa*) (EC 3.6.3.14)	190	−1·76 (0·35)	Oxidative phosphorylation
CN799346	18376041	ATP synthase epsilon chain (*N. crassa*) (EC 3.6.3.14)	193	2·55 (0·24)	Oxidative phosphorylation
CN799921	543870	ATP synthase protein 9 (*T. rubrum*) (EC 3.6.3.14)	349	2·22 (0·23)	Oxidative phosphorylation
CN794856	3023324	ATP synthase D chain (*Kluyveromyces lactis*) (EC 3.6.3.14)	267	1·19 (0·17)	Oxidative phosphorylation
Protein destination					
CN796287	16191634	Putative ubiquitin (*P. ostreatus*)	696	−1·32 (0·22)	Polyubiquitin
CN795232	9989034	Ubiquitin-conjugating enzyme (*Leishmania major*)	383	−1·46 (0·21)	Ubiquitin-conjugating enzyme E2
CN799723	6320718	Ubiquitin-like protein SMT3 (*S. cerevisiae*)	260	−1·62 (0·23)	Ubiquitin-like protein
CN795966	172260	Proteasome-related protein PRG1 (*S. cerevisiae*)	347	−2·39 (0·26)	20S proteasome beta-type subunit
CN794892	19115743	RER1-like protein-retention of ER proteins (*Sc. pombe*)	523	−1·36 (0·16)	
CN794605	16225946	Aspartyl aminopeptidase (*Coccidioides immitis*)	418	−3·23 (0·40)	Aminopeptidase
CN794779	14211580	Carboxypeptidase (*A. nidulans*)	540	−1·83 (0·21)	Carboxypeptidase
CN801309	1122233	Cuticle-degrading serine protease (*Arthrobotrys oligospora*)	225	2·48 (0·36)	Serine protease
CN800950	19171215	Alkaline serine protease (*Verticillium chlamydosporium*)	296	3·08 (0·24)	Serine protease
Protein synthesis					
CN796744	15226676	40S ribosomal protein S16 (*Arabidopsis thaliana*)	446	−2·72 (0·22)	Ribosomal protein
CN795526	6470325	Ribosomal protein L11 (*Sc. pombe*).	708	−2·52 (0·36)	Ribosomal protein
CN798832	5918018	L10A ribosomal protein (*C. albicans*)	515	2·40 (0·39)	Ribosomal protein
CN797582	19114836	Ribosomal protein l14 (*Sc. pombe*)	127	3·10 (0·35)	Ribosomal protein
CN794788	6006399	Ribosomal protein S28 (*A. niger*)	665	3·94 (0·30)	Ribosomal protein
CN794919	14318557	Ribosomal protein L29 (*S. cerevisiae*)	231	2·58 (0·23)	Ribosomal protein
CN795930	6325114	60S ribosomal protein L33-A (*S. cerevisiae*)	438	2·08 (0·18)	Ribosomal protein
CN794817	313261	Translation elongation factor eEF1 TEF1 (*S. cerevisiae*)	689	−1·71 (0·15)	Translation elongation factor
CN794828	19112922	Initiation factor eif-5a (*Sc. pombe*)	546	−1·33 (0·15)	Translation initiation factor
Transcription					
CN800324	122046	Histone H2B (*E. nidulans*)	526	−4·55 (0·27)	Histone
CN794916	17644129	Histone H2A (*N. crassa*)	498	−3·24 (0·19)	Histone
CN795244	122080	Histone H3 (*E. nidulans*)	625	−1·43 (0·16)	Histone
CN794487	122099	Histone H4 (*N. crassa*)	506	−2·92 (0·15)	Histone
CN795810	19067884	Hyperosmotic protein 21 (*Salmo salar*)	540	−2·83 (0·23)	Ring box protein
CN797125	1351369	Meiotic mRNA stability protein kinase UME5/SRP10 (*S. cerevisiae*)	261	−1·17 (0·18)	Component of RNA polymerase II holoenzyme

Table 4. cont.

EST clone	GenBank ID	GenBank homologue description	Score	Fold change	Function
CN799242	128183	Nonhistone chromosomal protein NHP6B (*S. cerevisiae*)	339	−1·80 (0·17)	HMG box protein
Unclassified protein					
CN794863	2133240	Immunoreactive spherule cell wall protein (*Co. immitis*)	194	−3·52 (0·17)	
CN800713	14029698	GAS1 (MAS3) protein (*Ma. grisea*)	311	1·77 (0·19)	
CN799200	15638627	Capsule-associated protein-like protein (*U. maydis*)	308	2·88 (0·30)	

microarray −2·03 and by real-time RT-PCR −0·25) was a *rho1* homologue, and the EST clone CN795728 (fold change on microarray −1·32) was a *rac1* homologue (Fig. 5). The *rho1* homologue of *M. haptotylum* (AY635989) encodes a protein of 193 aa. The deduced peptide rho1 contains all the signatures that have been found to be characteristic of the small G protein superfamily, including five GTP-binding and GTP hydrolysis domains (G1, G2, G3, G4 and G5) as well as the C-terminal prenylation site for rho GTPases (CXXL, C indicates cystein, X any amino acid and L leucin) involved in membrane associations (Bourne *et al.*, 1991; Finegold *et al.*, 1991). Rho1 of *M. haptotylum* shows high sequence similarity to the rho1 protein of other fungi, including that from *Aspergillus nidulans* (89 % aa identity), *Schizosaccharomyces pombe* (79 %), *S. cerevisiae* (76 %) and *Cryptococcus neoformans* (76 %).

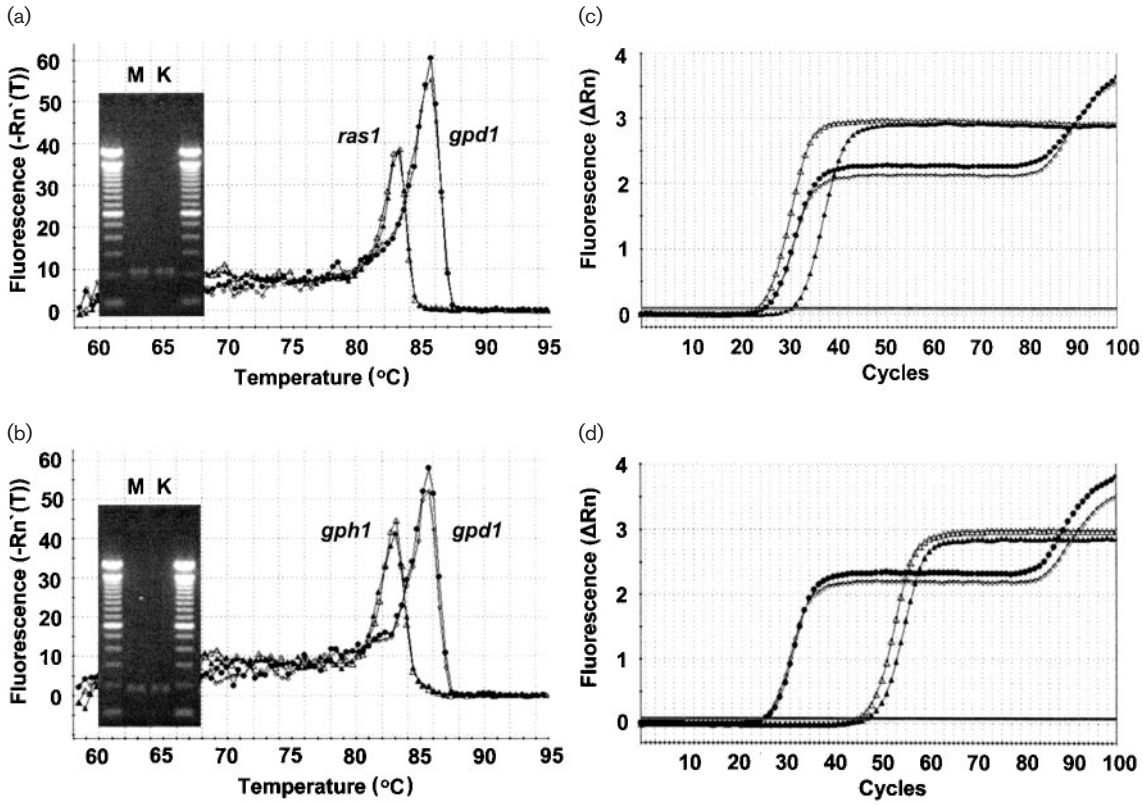

Fig. 3. Relative quantification of transcript levels for *ras1* and *gph1* in mycelium and knob tissue by the use of real-time quantitative RT-PCR. Dissociation curve profiles (temperature versus fluorescence) and amplification plots (cycle number versus fluorescence) are shown for *ras1* [(a) and (b)] and *gph1* [(c) and (d)], together with *gpd1*, which served as internal control for normalization (for gene identities, cf. Table 1). ●, amplification of *gpd1* in mycelium samples; ○, amplification of *gpd1* in knob samples; △, target genes (*ras1* or *gph1*) in knob samples; ▲, target genes in mycelium samples. Shown are mean values of triplicate cDNA samples made from aRNA. The dissociation curves show that the T_m values of the PCR products formed in the mycelium and knob samples were overlapping, indicating the specificity of the primers used. The inserted gels show that the length of the target amplicons was identical in the two tissues (M, mycelium; K, knob). In the amplification plots, the curves for the amplification of *gdp1* in the knob and mycelium samples were overlapping in the log-linear phase, indicating similar levels of *gpd1* transcripts in the two samples.

Table 5. Validation of microarray data

Presented are the mean fold change values ($\log_2$) of expression levels in knobs (trap cells) compared to mycelium. Three independent biological replicates were analysed. The same source of aRNA was used for both the cDNA microarray hybridization and the real-time RT-PCR quantification.

Gene	EST	Description	cDNA microarray	Real-time PCR	Validation
rho1	CN800249	GTPase Rho1	Down (−2·03)	Down (−0·25)	Yes
ras1	CN799032	Ras homologue	Up (+2·67)	Up (+2·19)	Yes
gks1	CN800713	GAS1 (MAS3) protein	Up (+1·77)	Up (+2·63)	Yes
gph1	CN795977	Glycogen phosphorylase	Up (+4·36)	Up (+2·89)	Yes
–	CN801309	Cuticle-degrading serine protease	Up (+2·48)	Up (+2·59)	Yes
–	CN800950	Alkaline serine protease	Up (+3·08)	Up (+2·79)	Yes
–	CN800820	Trichothecene 3-*O*-acetyltransferase	Up (+2·02)	Up (+2·17)	Yes
–	CN794863	Immunoreactive spherule cell wall protein	Down (−3·52)	Down (−0·27)	Yes

The switching between the GDP-bound inactive and the GTP-bound active form of rho GTPases is controlled by guanine nucleotide exchange factors (GEF) and GDP dissociation inhibitors (Koch *et al.*, 1997). A homologue of the yeast GDP dissociation inhibitor with activity toward rho1 (rdi1) was identified among the genes being down regulated in the knobs (EST clone CN794767, fold change on microarray −2·01). The *rdi1* gene of *M. haptotylum* (AY635991) is predicted to encode a 196 aa polypeptide, and displays 45 % aa identity to rdi1 of *S. cerevisiae.*

The *rac1* homologue of *M. haptotylum* (AY635990) was predicted to encode a peptide of 194 aa, and contained the conservative G1, G2, G3, G4 and G5 domains as well as the prenylation site (CXXL) characteristic of rac1 proteins (Bourne *et al.*, 1991). The protein showed 80 % sequence identity (aa) to a rac GTPase (*cflB*) of the dimorphic ascomycete *Penicillium marneffei* (Boyce *et al.*, 2003). Interestingly, cflB co-localizes with actin at the tips of the vegetative hyphal cells. Deletion of *cflB* results in growth defects in vegetative hyphae, depolarization and inappropriate septation (Boyce *et al.*, 2003).

The *ras1* homologue of *M. haptotylum* (AY635992, fold change on microarray +2·67 and by real-time RT-PCR +2·19) was predicted to encode a 198 aa protein, and showed a significant sequence similarity to ras-like proteins from *Colletotrichum trifolii* (46 % identity), *A. nidulans* (46 %), *S. cerevisiae* (42 %) and *Ustilago maydis* (44 %), especially in regions corresponding to the GTP-binding and GTP hydrolysis domains (G1, G2, G3, G4 and G5) as well as the C-terminal prenylation site (CAAX, C indicates

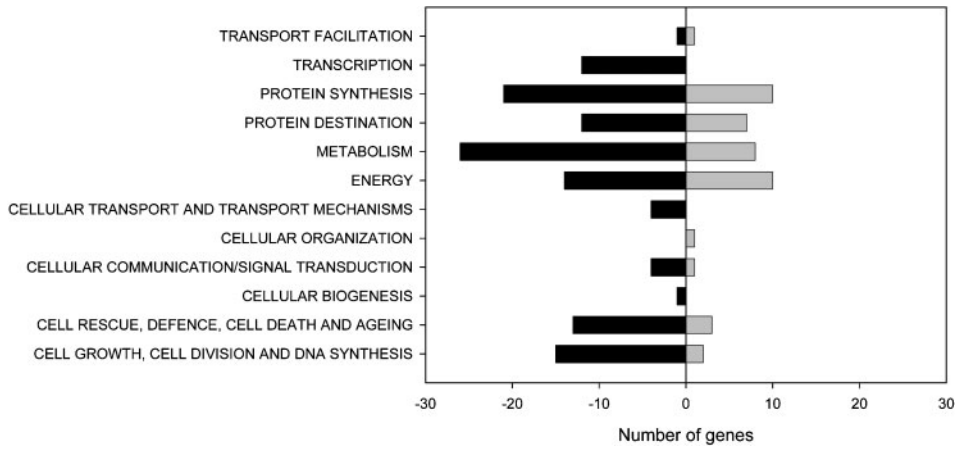

Fig. 4. Number of down- (left) and up- (right) regulated genes in knobs versus mycelium classified into various functional categories. Data are derived from cDNA microarray analysis using a significance level of $P<0{\cdot}05$ (cf. Table 3). The classification is according to MIPS (Mewes *et al.*, 2000), and annotation of putative functional roles is based on sequence similarity to information in the GenBank nr protein database. Of a total of 398 genes down-regulated, 123 genes were annotated into the displayed functional categories, whereas the remaining genes (not shown) were annotated either as unclassified proteins (179 genes) or as orphans with no sequence homology (96 genes). Among a total number of 259 upregulated genes, 43 were assigned functional annotation, whereas 119 clones were assigned as unclassified and 97 genes as orphans.

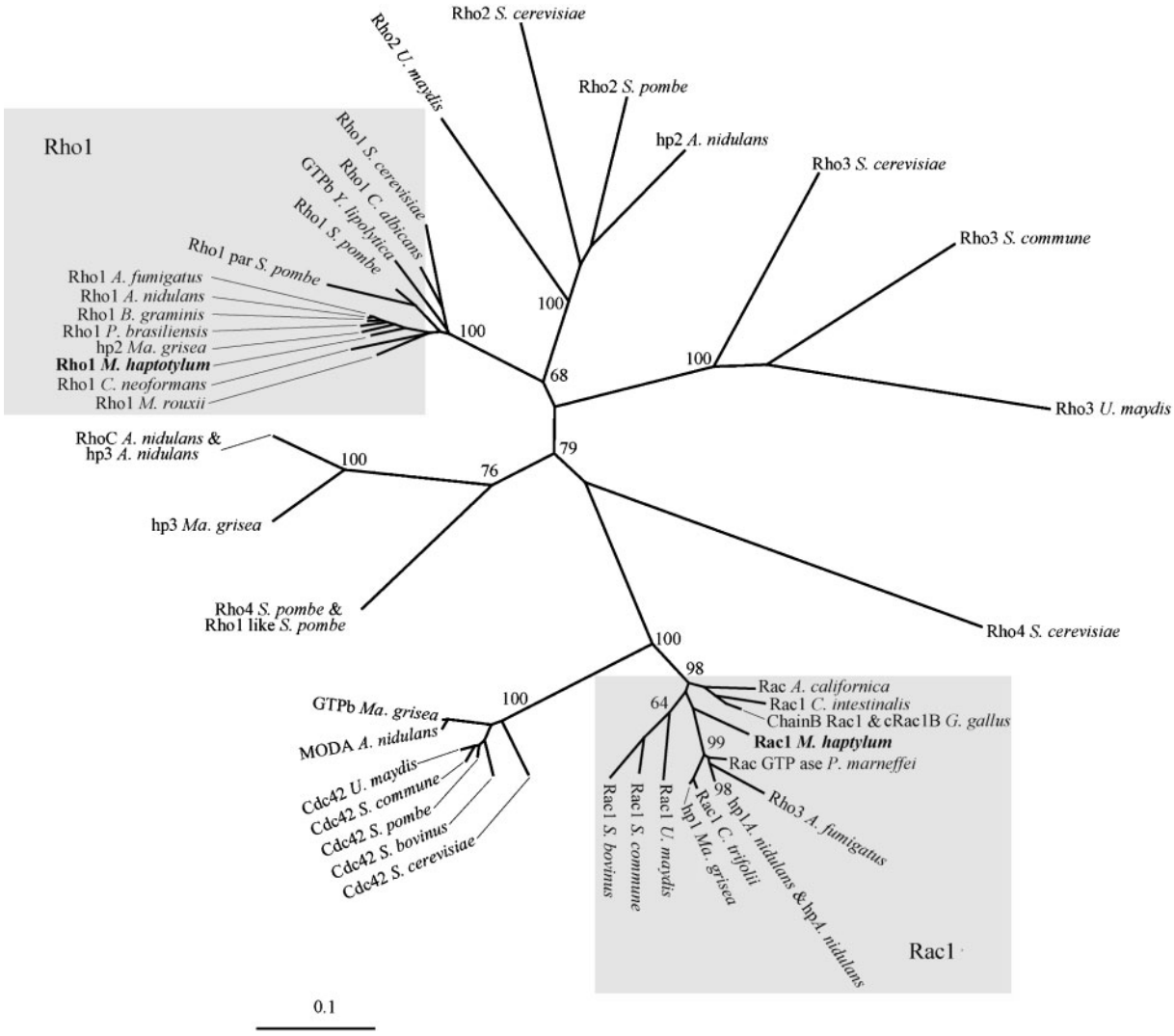

Fig. 5. Phylogeny of fungal rho and rac small GTPases. The tree (unrooted) was constructed using the neighbour-joining method (Saitou & Nei, 1987). The rho1 of *M. haptotylum* clusters with the rho1 sequences from other fungi (boxed), and the rac1 of *M. haptotylum* is found on a branch containing rac1 sequences from other fungi. The GenBank GI alt. accession numbers for sequences are (anticlockwise starting with rho1): 6325423, 3334315, 9887220, 19115402, 19114543, 113432032, 41017789, 15293426, 37549267, 38111209, 51556844 (AY635989), 3641690, 19909069. Rho3/Rho4: 41017782 (*A. nidulans*), 40743905, 38103445, 41017712, 19114956. Cdc42: 8132884, 6691157, 21667044, 15072535, 19114448, 7188786, 1071931. Rac1: 7188824 (*Saccharomyces bovinus*), 37181081, 21667516, 38100351, 32892148, 40741595, 40741595, 13432036, 25992183, 51556846 (AY635990), 3184512, 113096548, 30962119, 30385200. Rho4: 481101 (*S. cerevisiae*). Rho3: 21667514 (*U. maydis*), 12230529, 1071926. Rho2: 40741841 (*A. nidulans*), 19115481, 1301991 and 21667512.

cysteine, A an aliphatic amino acid and X any amino acid) (Bourne *et al.*, 1991).

Comparison of gene expression in knobs and appressoria

There are several similarities in the structure and function of knobs of nematode-trapping fungi and those of appressoria formed by plant-pathogenic fungi. Like a knob, an appressorium is a specialized infection structure, which develops as a spherical cell at the tip of a hypha (germ tube). Both structures contain an adhesive layer on the outside, which binds to the surface of the host. Furthermore, both appressoria and knobs form a hypha that penetrates the host using a combination of physical force and extracellular enzymic activities (Tucker & Talbot, 2001). Comparison of data from the transcriptional profiling of knobs with that of appressoria in *Magnaporthe grisea* and *Blumeria*

Table 6. Comparison of genes differentially regulated in knobs or appressoria in *M. haptotylum*, *Ma. grisea* and *B. graminis*, compared to mycelium, conidia and germinating conidia, respectively

For *M. haptotylum* knob, expression levels in knobs (trap cells) compared to mycelium are shown (Table 4). For *Ma. grisea* appressoria, expression levels in appressorium compared to differentiating conidia are shown (Takano *et al.*, 2004). For *B. graminis* appressoria, a comparison of appressoria and germinated conidia calculated from serial analysis of gene expression (SAGE) is shown (Thomas *et al.*, 2002).

Gene*	Knob *M. haptotylum*	Appressoria *Ma. grisea*	Appressoria *B. graminis*
gEgh16	Up	Up	Up
Chitinase	Down	Down	Up
Serine protease (subtilisin)	Up	Up	Up
Proteasome subunit	Down	Up	–
Cyclophilin	Up/down	–	Down
Peptidyl-prolyl cis-trans isomerase	Up	Up	Up
Metallothionein	Down	Up	Up
Thioredoxin	Down	–	Up
Translation elongation factor (TEF1)	Down	Down	–
14-3-3-like protein (BMH2)	Down	Up	Down
Triose phosphate isomerase	Down	–	Up
Glyceraldehyde-3-phosphate dehydrogenase	Down	Down	Down
Enolase	Down	–	Down
Fructose bisphosphate aldolase	Down	–	Down
GTPase Rho1	Down	–	Down
Profilin	Down	–	Down

*GenBank accession numbers. gEgh16: *M. haptotylum*, CN800713; *Ma. grisea*, AI069194; *B. graminis*, AW792327, AW789153 and AW789327. Chitinase: *M. haptotylum*, CN797557; *Ma. grisea*, BM872118; *B. graminis*, AW788278 and AW792426. Serine protease (subtilisin): *M. haptotylum*, CN801309, CN800950, CN799628 and CN800136; *Ma. grisea*, AI068925; *B. graminis*, AW790864. Proteasome subunit: *M. haptotylum*, CN795966; *Ma. grisea*, BG810344. Cyclophilin: *M. haptotylum*, CN800033, CN800928, CN800871, CN797293, CN794624 and CN795925; *B. graminis*, AW790370. Peptidyl-prolyl cis-trans isomerase: *M. haptotylum*, CN799935; *Ma. grisea*, AI068531; *B. graminis*, AW791080. Metallothionein: *M. haptotylum*, CN796974; *Ma. grisea*, CA408249; *B. graminis*, AW791711. Thioredoxin: *M. haptotylum*, CN795223; *B. graminis*, AW790498. Translation elongation factor TEF1: *M. haptotylum*, CN794817, CN795053, CN796490, CN795963 and CN799878; *Ma. grisea*, AI068515. 14-3-3-like protein (BMH2): *M. haptotylum*, CN795822; *Ma. grisea*, AI069388; *B. graminis*, AW790531. Triose phosphate isomerase: *M. haptotylum*, CN800416; *B. graminis*, AW789921. Glyceraldehyde-3-phosphate dehydrogenase: *M. haptotylum*, CN796696; *Ma. grisea*, CA408224; *B. graminis*: AW788939. Enolase: *M. haptotylum*, CN797534; *B. graminis*, AW789656. Fructose bisphosphate aldolase: *M. haptotylum*, CN797352; *B. graminis*, AW790281. GTPase Rho1: *M. haptotylum*, CN800249; *B. graminis*, AW792640. Profilin: *M. haptotylum*, CN795458; *B. graminis*: AW792002.

graminis shows that there are also many similarities in the patterns of gene regulation in the infection structures of nematode-trapping and plant-parasitic fungi (Table 6).

Among the commonly regulated genes was a gene (EST clone CN800713, fold change on microarray +1·77 and by real-time RT-PCR +2·63) encoding a polypeptide with sequence similarity to gas1 (previously mas3) of the rice blast fungus *Ma. grisea*. The gas1 protein is specifically expressed in appressoria and belongs to a large family that in *B. graminis* includes the gEgh16 and gEgh16H proteins (Xue *et al.*, 2002; Justesen *et al.*, 1996; Grell *et al.*, 2003). From complete cDNA sequencing of the gas1 homologue of *M. haptotylum* (AY635993), a 257 aa polypeptide could be predicted and the gene was designated *gks1* (gEgh16 homologue expressed in knob stage). The gks1 polypeptide displayed a 40 % aa sequence identity to gas1 of *Ma. grisea*. A phylogenetic analysis showed that gks1 and gas1 are found in a clade with high support (bootstrap value 73) containing members of the gEgh16 subfamily of the gEgh16/gEgh16H proteins (Fig. 6). We also identified another member of the gEgh16/gEgh16H family among the EST sequences of *M. haptotylum*. This EST clone (CN796958) was not differentially expressed, as revealed by the microarray analysis. The predicted polypeptide of CN796958 displayed significant sequence identity with gas2 (mas1) of *Ma. grisea* (43 % aa identity). CN796958 and gas2 were found in a well-supported clade (bootstrap value 100) containing members of the gEgh16H subfamily of the gEgh16/gEgh16H proteins (Fig. 6). The *gas1* and *gas2* genes showed a similar pattern of regulation in *Ma. grisea*, and both genes were expressed under appressorium formation (Xue *et al.*, 2002). In *B. graminis*, members of the gEgh16/gEgh16H family displayed different patterns of regulation, and several of them were differentially expressed during the development of the appressoria (Grell *et al.*, 2003; Thomas *et al.*, 2002).

Genes involved in stress and defence responses are one of the largest classes of genes that are differentially expressed

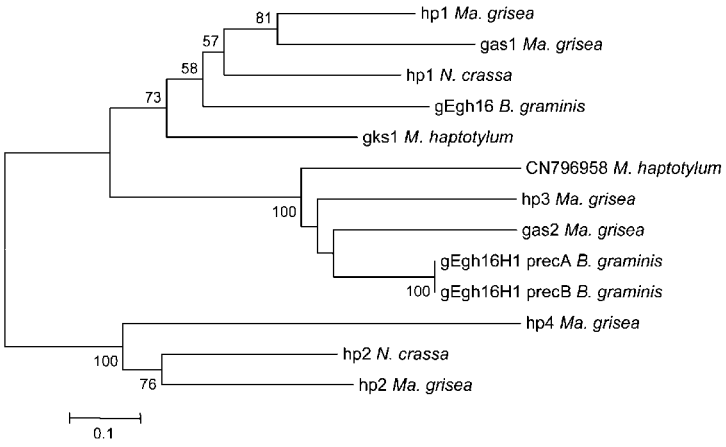

Fig. 6. Phylogeny of fungal gEgh16/gEgh16H proteins. The tree (unrooted) was constructed using the neighbour-joining method (Saitou & Nei, 1987). Gks1 of *M. haptotylum* is found in a clade with the Gas1 of *Ma. grisea* and gEgh16 of *B. graminis*. A second Gas homologue of *M. haptotylum* is found in a clade containing the Gas2 of *Ma. grisea* and the gEgh16H1 genes of *B. graminis*. Genbank GI alt. accession numbers of the protein/nucleotide sequences from top to bottom are: 38107786 (hp1 *Ma. grisea*), 14029698, 32416532, 2133327, 51556852 (AY635993), CN796958, 38108209, 8347747, 13430279, 13430281, 38102500, 32411089 and 38102176.

during appressoria formation in *Ma. grisea* (Rauyaree *et al.*, 2004). Several such genes, including cyclophilins, peptidyl-prolyl cis-trans isomerases, metallothionein and thioredoxins, appear to be differentially expressed in the knobs of *M. haptotylum* (Tables 4 and 6). Cyclophilins are a conserved family of proteins that direct and accelerate protein folding by a peptidyl cis-trans isomerase activity. In *Ma. grisea*, cyclophilin1 mutants showed reduced virulence and were impaired in penetration peg formation and generation of appressorium turgor pressure (Viaud *et al.*, 2002). In *B. graminis*, a putative metallothionein gene was identified as having high transcriptional levels in both conidia and appressoria (Thomas *et al.*, 2002). A transcript (EST clone CN796974) displaying sequence similarity to fungal Cu-binding metallothionein was differentially expressed in knobs (fold change on microarray −3·8).

A number of genes involved in protein synthesis (such as homologues for ribosomal proteins and translation elongation factor), protein destination and degradation (such as homologues for ubiquitin, ubiquitin-conjugating enzyme and proteasome components) were differentially expressed in both knobs and appressorium (Tables 4 and 5). This suggests that development of the infection structures in both nematode-trapping and plant-pathogenic fungi is associated with an extensive synthesis and turnover of proteins (McCafferty & Talbot, 1998).

Glycogen and carbon metabolism

One of the most upregulated genes in the knobs was a gene (EST clone CN795977, fold change on microarray +4·36 and by real-time RT-PCR +2·49) showing similarity to glycogen phosphorylases (α-D-glucosyltransferase, EC 2.4.1.1). Complete sequence analysis of this cDNA clone (AY635994) revealed an ORF corresponding to a putative protein of 874 aa. The entire aa sequence of *M. haptotylum* phosphorylase (gph1) showed 59 % identity to the *S. cerevisiae* glycogen phosphorylase (Hwang & Fletterick, 1986) and 45 % identity to rabbit muscle glycogen phosphorylase (Nakano *et al.*, 1986). Multiple sequence alignment showed extensive sequence similarity in regions corresponding to the catalytic domain and C-terminal domain, while the regulatory or N-terminal domain appeared more variable (Supplementary Fig. S1a with the online version of this paper at http://mic.sgmjournals.org/).

Extensive biochemical and crystallographic studies of the rabbit glycogen phosphorylase have identified seven functional domains, including the active site, the glycogen storage site, the purine nucleoside inhibitor site, the cofactor (pyridoxal phosphate, PLP) binding site, the phosphorylation site, the AMP binding site, and the subunit dimerization point (Hwang & Fletterick, 1986). Four of these regions, the active site (labelled as *g* in Supplementary Fig. S1a) (1 substitution in 17 residues), the glycogen storage site (*s*) (2 substitutions in 8 residues), the purine nucleoside inhibitor site (*c*) (1 substitution in 4 residues) and the PLP binding site (*v*) (1 substitution in 8 residues) were highly conserved between the *M. haptotylum* and the rabbit enzymes. The N-terminal regulatory domain is important for allosteric regulation, and carries the regulatory phosphorylation site (*p*) of glycogen phosphorylases. In the rabbit enzyme, this domain is located within the first 80 aa of the N-terminal region (Palm *et al.*, 1985). Apart from the phosphorylation site, the N-terminal region of the phosphorylases contains the AMP binding site (*a*) and the subunit dimerization point (*d*). Considerable amino acid sequence identity (76 %) was observed in the N-terminal region of the yeast and the *M. haptotylum* phosphorylases, whereas the rabbit enzyme appeared more divergent. The yeast enzyme is phosphorylated at a specific threonine residue (Thr-19) by a phosphorylase kinase, and additionally by a cyclic AMP-dependent protein kinase (Hwang & Fletterick, 1986). The Thr-19 residue, as well as adjacent residues, were well conserved between the yeast and the *M. haptotylum* phosphorylase.

A phylogenetic analysis of phosphorylases from different organisms showed that the gph1 protein of *M. haptotylum*

was found in a clade with high bootstrap support (value 100), containing sequences from several ascomycetes, including a putative glycogen phosphorylase of the human pathogen *Aspergillus fumigatus* and hypothetical proteins from the rice blast fungus *Ma. grisea* (Supplementary Fig. S1b).

Glycogen is an important storage carbohydrate in eukaryotes, and glycogen phosphorylase catalyses and regulates the degradation of glycogen to glucose 1-phosphate (Fletterick & Madsen, 1980). Glycogen is accumulated in incipient appressoria of *Ma. grisea*, but is rapidly degraded before generating the turgor pressure needed for penetration of the plant cuticle (Thines *et al.*, 2000). The turgor pressure results from a rapid accumulation of glycerol, and there are different lines of evidence suggesting that the production of glycerol is achieved by the mobilization of energy reserves, such as glycogen and neutral lipids (Thines *et al.*, 2000). The glucose 1-phosphate generated from the degradation of glycogen is metabolized in the glycolytic pathway. Glycerol can be synthesized from several of the intermediates in this pathway. Unfortunately, we did not identify any putative homologues in the *M. haptotylum* EST database of enzymes known to be involved in these pathways in *S. cerevisiae* (see references in Thines *et al.*, 2000). Thus, whether or not the genes encoding these enzymes are regulated in the knobs is not known. Alternatively, the breakdown of glycogen can lead to the formation of pyruvate and the production of energy (ATP). However, this part of the pathway does not appear to be upregulated in the knobs. The homologues of enzymes in this part of the glycolytic pathway that were identified, including glyceraldehyde-3-phosphate dehydrogenase (CN796696) and enolase (CN797534), were downregulated in the knobs compared to mycelium (fold change on microarray, −1·59 and −2·07, respectively) (Table 4). Studies have also shown that the degradation of glycogen in *Ma. grisea* is under the control of the cAMP-dependent protein kinase A (PKA) and MAP kinase (MAPK) pathway (Thines *et al.*, 2000). In yeast, phosphorylase activity is regulated by reversible phosphorylation/dephosphorylation mediated by the cAMP regulatory cascade (Hwang & Fletterick, 1986). The fact that the N-terminal domain, which contains regions important for allosteric regulation and carries the regulatory phosphorylation site (Thr-19), was well conserved between the yeast and *M. haptotylum* enzymes, suggests that the two phosphorylases may be regulated by similar mechanisms.

Peroxisomal associated proteins

The presence of numerous dense bodies, which are related to peroxisomes, is typical for traps of nematode predatory fungi (Dijksterhuis *et al.*, 1994). Interestingly, one transcript (CN796915, fold change on microarray +0·72) with similarity to a peroxisomal membrane protein in *Candida boidiini* (PMP30B) was significantly upregulated. There is a homologue in *S. cerevisiae* (PMP27, PEX11, peroxin) which is involved in peroxisomal proliferation. Deletion of PMP27 generates a phenotype containing a few large peroxisomes, as if peroxisomal fission was inhibited (Erdmann & Blobel, 1995). Other genes that were regulated in knobs and displayed sequence similarities to genes encoding proteins associated with peroxisomes included homologues to D-amino acid oxidase (CN796046, fold change on microarray −3·09), 2,4-dienoyl-CoA reductase (SPS19, *S. cerevisiae*) (CN797536, fold change on microarray +1·05), and fatty acid transporter and very long-chain fatty acyl-CoA synthetase (FAT1, *S. cerevisiae*) (CN796073, fold change on microarray −2·30).

ACKNOWLEDGEMENTS

This paper is dedicated to Professor Birgit Nordbring-Hertz to honour her pioneer work on the morphology and physiology of nematode-trapping fungi. Custom microarrays were produced at the SWEGENE DNA Microarray Resource Center at the BioMedical Center B10 in Lund, and DNA sequencing was performed at the SWEGENE Center of Genomic Ecology at the Ecology Building in Lund, supported by the Knut and Alice Wallenberg Foundation through the SWEGENE consortium. The study was supported by grants from the Swedish Natural Science Research Council. We wish to thank Björn Canbäck and Balaji Rajashekar for their kind help on bioinformatics issues.

REFERENCES

Ahren, D., Troein, C., Johansson, T. & Tunlid, A. (2004). PHOREST: a web-based tool for comparative analyses of EST data. *Mol Ecol Notes* **4**, 311–314.

Ahrén, D., Ursing, B. M. & Tunlid, A. (1998). Phylogeny of nematode-trapping fungi based on 18S rDNA sequences. *FEMS Microbiol Lett* **158**, 179–184.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

Ayscough, K. R. (1998). *In vivo* functions of actin-binding proteins. *Curr Opin Cell Biol* **10**, 102–111.

Barron, G. L. (1977). *The Nematode-Destroying Fungi.* Guelph, Canada: Lancester Press.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. (2004). GenBank: update. *Nucleic Acids Res* **32**, 23–26.

Bourne, H. R., Sanders, D. A. & McCormick, F. (1991). The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**, 117–127.

Boyce, K. J., Hynes, M. J. & Andrianopoulos, A. (2003). Control of morphogenesis and actin localization by the *Penicillium marneffei* RAC homolog. *J Cell Sci* **116**, 1249–1260.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans. Genetics* **77**, 71–94.

Dijksterhuis, J., Veenhuis, M., Harder, W. & Nordbring-Hertz, B. (1994). Nematophagous fungi: physiological aspects and structure-function relationships. *Adv Microb Physiol* **36**, 111–143.

Emmons, S. W., Klass, M. R. & Hirsh, D. (1979). Analysis of the constancy of DNA sequences during development and evolution of the nematode *Caenorhabditis elegans. Proc Natl Acad Sci U S A* **76**, 1333–1337.

Erdmann, R. & Blobel, G. (1995). Giant peroxisomes in oleic acid-induced *Saccharomyces cerevisiae* lacking the peroxisomal membrane protein Pmp27p. *J Cell Biol* **128**, 509–523.

Ewing, B., Hillier, L., Wendl, M. C. & Green, P. (1998). Base-calling of automated sequencer traces using *Phred*. I. Accuracy assessment. *Genome Res* **8**, 175–185.

Finegold, A. A., Johnson, D. I., Farnsworth, C. C., Gelb, M. H., Judd, S. R., Glomset, J. A. & Tamanoi, F. (1991). Protein geranylgeranyltransferase of *Saccharomyces cerevisiae* is specific for Cys-Xaa-Xaa-Leu motif proteins and requires the CDC43 gene product but not the DPR1 gene product. *Proc Natl Acad Sci U S A* **88**, 4448–4452.

Fletterick, R. J. & Madsen, N. B. (1980). The structures and related functions of phosphorylase *a*. *Annu Rev Biochem* **49**, 31–61.

Friman, E. (1993). Isolation of trap cells from the nematode-trapping fungus *Dactylaria candida*. *Exp Mycol* **17**, 368–370.

Grell, M. N., Mouritzen, P. & Giese, H. (2003). A *Blumeria graminis* gene family encoding proteins with a C-terminal variable region with homologues in pathogenic fungi. *Gene* **311**, 181–192.

Huang, X. (1992). A contig assembly program based on sensitive detection of fragment overlaps. *Genomics* **14**, 18–25.

Hwang, P. K. & Fletterick, R. J. (1986). Convergent and divergent evolution of regulatory sites in eukaryotic phosphorylases. *Nature* **324**, 80–84.

Johansson, T., Le Quéré, A., Ahrén, D., Söderström, B., Erlandsson, R., Lundeberg, J., Uhlén, M. & Tunlid, A. (2004). Transcriptional responses of *Paxillus involutus* and *Betula pendula* during formation of ectomycorrhizal root tissue. *Mol Plant–Microbe Interact* **17**, 202–215.

Justesen, A., Somerville, S., Christiansen, S. & Giese, H. (1996). Isolation and characterization of two novel genes expressed in germinating conidia of the obligate biotroph *Erysiphe graminis* f.sp. *hordei*. *Gene* **170**, 131–135.

Kerr, M. K. & Churchill, G. A. (2001). Experimental design for gene expression microarrays. *Biostatistics* **2**, 183–201.

Koch, G., Tanaka, K., Masuda, T., Yamochi, W., Nonaka, H. & Takai, Y. (1997). Association of the Rho family small GTP-binding proteins with Rho GDP dissociation inhibitor (Rho GDI) in *Saccharomyces cerevisiae*. *Oncogene* **15**, 417–422.

Kumar, S., Tamura, K., Jakobsen, I. B. & Nei, M. (2001). MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**, 1244–1245.

Lappalainen, P. & Drubin, D. G. (1997). Cofilin promotes rapid actin filament turnover *in vivo*. *Nature* **388**, 78–82.

Larsen, M. (2000). Prospects for controlling animal parasitic nematodes by predacious microfungi. *Parasitology* **120**, 121–131.

Livak, K. J. & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* **25**, 402–408.

McCafferty, H. R. & Talbot, N. J. (1998). Identification of three ubiquitin genes of the rice blast fungus *Magnaporthe grisea*, one of which is highly expressed during initial stages of plant colonisation. *Curr Genet* **33**, 352–361.

Mewes, H. W., Frishman, D., Gruber, C. & 7 other authors (2000). MIPS: a database for genomes and protein sequences. *Nucleic Acids Res* **28**, 37–40.

Nakano, K., Hwang, P. K. & Fletterick, R. J. (1986). Complete cDNA sequence for rabbit muscle glycogen phosphorylase. *FEBS Lett* **204**, 283–287.

Nordbring-Hertz, B., Jansson, H.-B., Friman, E. & 5 other authors (1995). *Nematophagous fungi*. Film No C 1851. Göttingen, Germany: Institut für den Wissenschaftlichen Film.

Palm, D., Goerl, R. & Burger, K. J. (1985). Evolution of catalytic and regulatory sites in phosphorylases. *Nature* **313**, 500–502.

Pearson, W. R. (1994). Using the FASTA program to search protein and DNA sequence databases. *Methods Mol Biol* **24**, 307–331.

Pruyne, D. & Bretscher, A. (2000a). Polarization of cell growth in yeast. I. Establishment and maintenance of polarity states. *J Cell Sci* **113**, 365–375.

Pruyne, D. & Bretscher, A. (2000b). Polarization of cell growth in yeast. II. The role of the cortical actin cytoskeleton. *J Cell Sci* **113**, 571–585.

Rauyaree, P., Choi, W., Fang, E., Blackmon, B. & Dean, R. A. (2004). Genes expressed during early stages of rice infection with the rice blast fungus *Magnaporthe grisea*. *Mol Plant Pathol* **2**, 347–354.

Rozen, S. & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. In *Bioinformatics Methods and Protocols*, pp. 365–386. Edited by S. Krawetz & S. Misener. Totowa, NJ: Humana Press.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406.

Stiekema, W. J., Heidekamp, F., Dirkse, W. G., van Beckum, J., de Haan, P., ten Bosch, C. & Louwerse, J. D. (1988). Molecular cloning and analysis of four potato tuber mRNA. *Plant Mol Biol* **11**, 255–269.

Takano, Y., Choi, W., Mitchell, T. K., Okuno, T. & Dean, R. A. (2004). Large scale parallel analysis of gene expression during infection-related morphogenesis of *Magnaporthe grisea*. *Mol Plant Pathol* **4**, 337–346.

Thines, E., Weber, R. W. & Talbot, N. J. (2000). MAP kinase and protein kinase A-dependent mobilization of triacylglycerol and glycogen during appressorium turgor generation by *Magnaporthe grisea*. *Plant Cell* **12**, 1703–1718.

Thomas, S. W., Glaring, M. A., Rasmussen, S. W., Kinane, J. T. & Oliver, R. P. (2002). Transcript profiling in the barley mildew pathogen *Blumeria graminis* by serial analysis of gene expression (SAGE). *Mol Plant–Microbe Interact* **15**, 847–856.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

Tucker, S. L. & Talbot, N. J. (2001). Surface attachment and pre-penetration stage development by plant pathogenic fungi. *Annu Rev Phytopathol* **39**, 385–417.

Tunlid, A., Johansson, T. & Nordbring-Hertz, B. (1991). Surface polymers of the nematode-trapping fungus *Arthrobotrys oligospora*. *J Gen Microbiol* **137**, 1231–1240.

Viaud, M. C., Balhadere, P. V. & Talbot, N. J. (2002). A *Magnaporthe grisea* cyclophilin acts as a virulence determinant during plant infection. *Plant Cell* **14**, 917–930.

Wolfinger, R. D., Gibson, G., Wolfinger, E. D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C. & Paules, R. S. (2001). Assessing gene significance from cDNA microarray expression data via mixed models. *J Comput Biol* **8**, 625–637.

Wu, T. D. (2001). Analysing gene expression data from DNA microarrays to identify candidate genes. *J Pathol* **195**, 53–65.

Xue, C., Park, G., Choi, W., Zheng, L., Dean, R. A. & Xu, J. R. (2002). Two novel fungal virulence genes specifically expressed in appressoria of the rice blast fungus. *Plant Cell* **14**, 2107–2119.