Microbiology (2003), **149**, 239–247 DOI 10.1099/mic.0.25761-0

Expressed sequence tag (EST) analysis of two subspecies of *Metarhizium anisopliae* reveals a plethora of secreted proteins with potential activity in insect hosts

Florian M. Freimoser, Steven Screen,† Savita Bagga, Gang Hu and Raymond J. St Leger

Department of Entomology, University of Maryland, 4112 Plant Sciences Building, College Park, MD, 20742, USA

Correspondence
Raymond J. St Leger
rl106@umail.umd.edu

Received 24 May 2002
Revised 1 October 2002
Accepted 15 October 2002

Expressed sequence tag (EST) libraries for *Metarhizium anisopliae*, the causative agent of green muscardine disease, were developed from the broad host-range pathogen *Metarhizium anisopliae* sf. *anisopliae* and the specific grasshopper pathogen, *M. anisopliae* sf. *acridum*. Approximately 1700 5′ end sequences from each subspecies were generated from cDNA libraries representing fungi grown under conditions that maximize secretion of cuticle-degrading enzymes. Both subspecies had ESTs for virtually all pathogenicity-related genes cloned to date from *M. anisopliae*, but many novel genes encoding potential virulence factors were also tagged. Enzymes with potential targets in the insect host included proteases, chitinases, phospholipases, lipases, esterases, phosphatases and enzymes producing toxic secondary metabolites. A diverse array of proteases composed 36 % of all *M. anisopliae* sf. *anisopliae* ESTs. Eighty percent of the ESTs that could be clustered into functional groups had significant matches ($E < 10^{-5}$) in other ascomycete fungi. These included genes reported to have specific roles in pathogens with plant or vertebrate hosts. Many of the remaining ESTs had their best BLAST match among animal, plant and bacterial sequences. These include genes with plant and microbial counterparts that produce potent antimicrobials. The abundance of transcripts discovered for different functional groups varied between the two subspecies of *M. anisopliae* in a manner consistent with ecological adaptations of the two pathogens. By hastening gene discovery this project has enhanced development of improved mycoinsecticides. In addition, the *M. anisopliae* ESTs represent a significant contribution to the extensive database of sequences from ascomycetes that are saprophytes or plant and vertebrate pathogens. Comparative analyses of these sequences is providing important information about the biology and evolutionary history of this clade.

INTRODUCTION

At least 90 genera and more than 700 species of fungi, dispersed in virtually every major taxonomic group except the higher basidiomycetes, have been identified as insect pathogens (Roberts & Humber, 1981). In terms of diversity they rival the plant pathogens, while there are comparatively few (about 40) fungal pathogens of warm-blooded animals (Rippon, 1988). Insect-pathogenic fungi such as *Metarhizium anisopliae* have been extensively studied as key regulatory factors in insect populations and as agents of biocontrol (Inglis *et al.*, 2001). However, attempts to discern the physiological determinants of infection processes and produce a rational plan for strain improvement have often been thwarted by the complexity of fungal responses to host-related signals. Consequently, side effects occurring in a selected or constructed strain are hard to predict and assess, and the full range of engineering possibilities cannot be exploited, due to lack of knowledge of inter-related regulatory and metabolic processes going on in the cell.

We need alternative strategies to assess genomes of *M. anisopliae* and its fellow insect pathogens on a large scale and thus ensure continued advances in the biology of these important pest control agents. We consequently adopted an EST strategy to assess pathogenicity determinants in

†Present address: Bioinformatics Lead Identification, The Monsanto Company, 800 N. Lindbergh Blvd, St Louis, MO 63167, USA.

Abbreviation: EST, expressed sequence tag.

The NCBI accession numbers for the sequences reported in this paper are listed in Tables available as supplementary data with the on-line version of this paper at http://mic.sgmjournals.org

M. anisopliae as 'whole systems' rather than as isolated parts. *M. anisopliae* is an obvious choice to initiate such studies. Although with the exception of proteases, only a few potential *M. anisopliae* virulence factors have been examined in detail (either biochemically or genetically) there are sufficient biochemical and molecular data to allow predictions on the likely role of virulence factors in killing an insect, overcoming the insect immune system and in facilitating fungal growth. Information at the same level is not available for other fungal pathogens of insects (St Leger & Screen, 2001). However, like *M. anisopliae* these fungi penetrate host cuticle directly, render host tissues suitable for consumption and overcome cell- and peptide-mediated components of the insect immune system. Broadly dispersed, anti-insect virulence mechanisms might, therefore, include toxins and hydrolytic enzymes capable of degrading host tissues and disabling anti-microbial peptides. Furthermore, we might expect pathogens to produce a broad array of antimicrobials to defend against opportunistic colonizers of the insect cadavers.

On a wider scale, *M. anisopliae* belongs to the clavicipitaceous pyrenomycetes and is closely related to genera containing such well-known animal and plant pathogens as *Fusarium*, *Trichoderma* and *Paecilomyces*. However, while it has been shown that these fungal pathogens share some enzymes involved in pathogenicity (Reddy *et al.*, 1996), there is little overlap between plant and insect pathogens at the genus level. It follows that the characteristics needed by fungi to successfully establish disease in plants must be fundamentally different in some ways from those needed to infect animals. It is an underlying assumption of our work that as well as providing a model for insect pathogens, identifying the nature and networking of the genes required during pathogenicity in *Metarhizium* spp. will enrich understanding of fungal pathogenesis in these other systems and delineate probable key virulence characters for pathogens of different host groups.

Current EST projects for related ascomycetes such as *Neurospora crassa* mostly employ libraries obtained from sporulation cultures and from mycelium growing in much more nutrient-rich (i.e., catabolite-repressing) conditions than would be met on a host surface (Nelson *et al.*, 1997). Such conditions repress expression of secreted proteins in *Metarhizium* spp., including known pathogenicity determinants (St Leger *et al.*, 1986b). Although EST projects are apparently random in nature, it is possible to target this type of project at genes involved in specific processes. It is most likely that the majority of parasitism-related genes encode secreted molecules, as these will be in most intimate contact with the host. Indeed, the only putative pathogenicity genes cloned from entomopathogenic fungi to date have encoded secreted molecules responsible for solubilizing host barriers, acquisition of nutrients and toxic effects against the host and competing microbes (St Leger & Bidochka, 1996). The libraries we employed in this study were made from fungi growing on insect cuticles, which induce production of great quantities of secreted proteins, including (candidate) pathogenicity genes (St Leger *et al.*, 1994). mRNAs encoding secreted products are therefore likely to be present in abundance in a representative cDNA library.

In this study we compare ESTs from two subspecies of *M. anisopliae* that have been widely employed as biological insecticides (Inglis *et al.*, 2001). Like many ascomycete fungi, *Metarhizium anisopliae* sf. *anisopliae* (ARSEF 2575) is a facultative saprophyte with both free-living (saprophytic) and pathogenic life stages. It is a cosmopolitan pathogen and has been reported from over 200 insect species (Samuels *et al.*, 1989). In contrast, subspecies *M. a.* sf. *acridum* (ARSEF 324) has a very limited host-range, being only known to attack orthopteran insects, e.g., grasshoppers (Inglis *et al.*, 2001). Reflecting this, *M. a.* sf. *acridum* is less plastic in its physiological responses and unlike *M. a.* sf. *anisopliae* requires host-related stimuli such as chitin or cuticle to germinate and produce infection structures (St Leger *et al.*, 1992). While *M. a.* sf. *anisopliae* produces multiple proteases, *M. a.* sf. *acridum* produces more types and greater amounts of chitinolytic enzymes (St Leger *et al.*, 1993a, 1996c), suggesting that a better estimate of the range of *M. anisopliae* pathogenicity determinants may be obtained by studying both subspecies.

METHODS

Strains and culture conditions. The fungal strains used in this study were *M. a.* sf. *anisopliae* (ARSEF 2575) and *M. a.* sf. *acridum* (ARSEF 324). These fungi were routinely grown at 27 °C either in liquid (SDB) or on solid (SDA) Sabouraud-glucose medium supplemented with 0·5 % yeast extract. For cDNA library construction, fungi were first grown for 30 h in liquid SDB broth. The cultures were then washed with sterile distilled water and the fungal biomass was transferred for 24 h to minimal medium (containing 0·1 %, w/v, KH_2PO_4, 0·05 % (w/v) $MgSO_4$ and 50 %, v/v, tap water) supplemented with 1 % (w/v) cockroach cuticle (for *M. a.* sf. *anisopliae*) or 1 % (w/v) chitin and 1 % (w/v) cockroach cuticle (for *M. a.* sf. *acridum*). Addition of chitin to growth medium is required to maximize protein secretion by *M. a.* sf. *acridum* and induce putative pathogenicity factors (St Leger *et al.*, 1992, 1993a, 1994).

Construction of cDNA libraries. Total RNA was extracted from frozen fungus using TRI Reagent as described by Joshi & St Leger (1999). The cDNA libraries were constructed in the unidirectional λZAP II vector (Stratagene) exploiting the *Eco*RI and *Xho*I restriction sites. The cDNA libraries were not normalized, i.e. no attempt was made to reduce the redundancy of highly expressed transcripts.

Plasmid isolation and DNA sequencing. Plasmid constructs were transformed in *Escherichia coli* TOP10 (Invitrogen). Individual transformants were picked, grown overnight in LB medium and plasmid DNA was isolated and purified using QIAprep Spin Miniprep Kits (Qiagen) following the company's protocols. cDNA inserts were sequenced from the 5′ end by employing the M13 primer and ABI chemicals on ABI 377 DNA sequencers (DNA Sequencing Facility, Center for Agricultural Biotechnology, University of Maryland, USA).

Sequence analysis. Vector sequences were removed by hand. Overlapping sequences were assembled into consensus sequences

(contigs) by using the program CAP3 (Huang & Madan, 1999). The program BLASTX (Altschul *et al.*, 1997) was used to search all ESTs against the non-redundant amino acid reference library (NCBI's 'nr' database) or against an amino acid database containing only fungal sequences. In general, similarities with *E*-values $< 10^{-5}$ were considered significant (Anderson & Brass, 1998). All sequences were submitted to NCBI. Accession numbers for *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum* begin with prefixes AJ and BQ, respectively. Accession numbers and results from BLAST searches are given in the supplementary data Tables available with the on-line version of this paper at http://mic.sgmjournals.org

RESULTS AND DISCUSSION

Characterization of the EST database

The mean insert size for the *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum* cDNA libraries were 1·0 kb (0·5–2 kb) and 1·5 kb (0·8–2·7 kb), respectively. Around 2000 clones were picked at random from each library and 1692 (*M. a.* sf. anisopliae) and 1727 (*M. a.* sf. *acridum*) sequences were obtained. After editing, the mean length of the single read sequences was 520 bp. Assembling the ESTs into contigs resulted in 881 unigenes (unique EST sequences, including singletons) for *M. a.* sf. *anisopliae* and 1211 unigenes for *M. a.* sf. *acridum*. Contigs that consisted of one sequence (singletons) composed 736/1692 and 1063/1727 sequences in *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum*, respectively. Database searching resulted in 491 (*M. a.* sf. *anisopliae*) and 448 (*M. a.* sf. *acridum*) unique ESTs with significant matches ($E \leqslant 10^{-5}$).

Comparative analyses indicate varied patterns of similarity to other ESTs

ESTs with *E* values of $\leqslant 10^{-5}$ were grouped into functional categories as outlined in Fig. 1 (for further details see the Tables available as supplementary data at http://mic.sgmjournals.org). Not surprisingly, about 80 % of the ESTs from both fungi that could be assigned to a functional category ($E \leqslant 10^{-5}$) had a fungal sequence as the best match. These hits were almost exclusively among ascomycete sequences; only 2 % (*M. a.* sf. *anisopliae*) or 1 % (*M. a.* sf. *acridum*) of the ESTs had their best match sequences from Basidiomycetes. About 8 % of the sequences from *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum* were most similar to sequences from animals or bacteria. A smaller number of EST sequences of *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum* showed most similarity to plant proteins, while very few ESTs had their closest counterpart in an archean or protistan sequence.

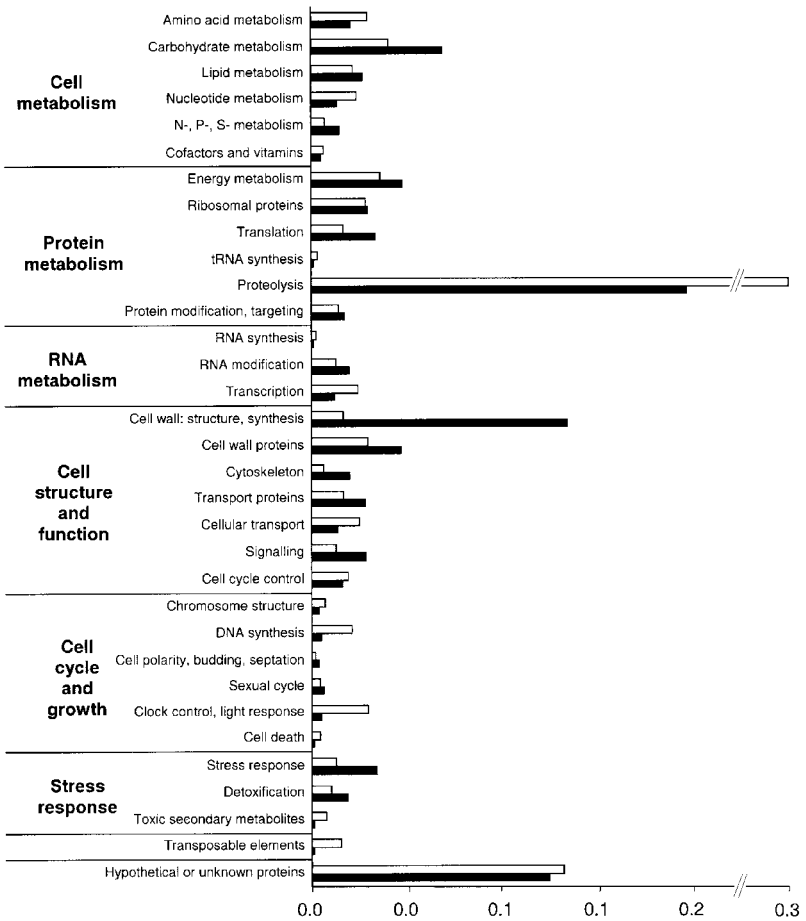

Fig. 1. Proportions of EST sequences with significant BLAST hits ($E < 10^{-5}$) from *M. a.* sf. *anisopliae* (white bars) and *M. a.* sf. *acridum* (black bars) assigned to the listed functional categories.

Putative identification of pathogenicity-related genes

Our main focus is aimed at cDNAs encoding proteins that may be involved in the fungus–insect interaction. A striking feature of both libraries is the high representation of secreted depolymerases that we had predicted from the choice of growth conditions for library construction. Maximizing protein secretion by *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum* was intended to mimic conditions occurring during infection and colonization of the host insect and increase the probability of cloning transcripts involved in pathogenicity. The success of this strategy was supported by expression of several EST sequences, including proteases (Pr1a, chymotrypsin, trypsins, carboxypeptidase), chitinases and a cAMP-dependent protein kinase catalytic subunit (AJ273794) that are produced *in vivo* and have been implicated in the events leading to formation of infection structures and pathogenicity (Joshi & St Leger, 1999; Screen & St Leger, 2000; St Leger *et al.*, 1996b, c).

It is difficult to determine the involvement in parasitism of ESTs without significant matches in databases. These genes require additional information regarding their function and likewise the genes with homologues of unknown function in other systems. Thus, the highly expressed AJ274161 (Table 1 and see Table A available as supplementary data at http://mic.sgmjournals.org) is similar ($E = 5 \times 10^{-24}$) to rASP f4, an antigen in the human pathogen *Aspergillus fumigatus* (Crameri & Blaser, 1996). Their function in pathogenicity is not evident. However, as insects provide a much more convenient model system than vertebrates we are in a position to begin to understand the roles of individual genes of unknown function via targeted disruption and possibly extrapolate these findings to vertebrate pathogens.

For other genes the function of putative homologues may be indicative of a role in fungal pathogenicity. Matches to secreted hydrolytic enzymes with substrates in insect hosts composed the most abundant classes of ESTs in both *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum*. These include proteases, chitinases, phospholipases, lipases, esterases and phosphatases. Furthermore, most of these contigs were redundant (consisting of two or more transcripts), allowing electronic Northerns to measure transcript abundance (Table 1). Proteases compose 36 and 20 % of the *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum* libraries, respectively (Fig. 1). This is consistent with high levels of expression of proteases reported previously (St Leger *et al.*, 1989). No other fungal species has been reported to have such a diverse array of proteases as *M. anisopliae* with many different subtilisins (8 in *M. a.* sf. *anisopliae* and 3 in *M. a.* sf. *acridum*), trypsins (AJ274329, AJ273017), a chymotrypsin (AJ273663), metalloproteases (AJ273481, BQ142817, BQ143246), aspartyl proteases (AJ274168, BQ143130), aminopeptidases (AJ273956, AJ274066), X-Pro-dipeptidyl aminopeptidase (AJ274266) and carboxypeptidases (AJ274102, AJ272919, BQ142584). Many of these activities have already been characterized biochemically in culture filtrates of strain 2575 (e.g. St Leger *et al.*, 1993a, b, 1994). A sequence similar ($E = 2 \times 10^{-12}$) to a calpain-like (cysteine) protease from *Aspergillus oryzae* was also detected in *M. a.* sf. *acridum* (BQ142792).

The comparative paucity of proteases is not likely to represent incomplete sampling in other fungi as the completed genomes of *Saccharomyces cerevisiae*, *Neurospora crassa* and *Aspergillus nidulans*, and extensive EST-collections from other ascomycetes, contain no chymotrypsins, often no trypsins and at most two or three subtilisin genes. Previous attempts to profile *M. a.* sf. *anisopliae* subtilisins using differential display and differential hybridization identified two genes, *Pr1*a and *Pr1*b (Joshi *et al.*, 1997; St Leger *et al.*, 1996a), and *Pr1e* (AJ251967) and *Pr1f* (AJ251967) (S. Bagga, S. Screen, F. M. Freimoser & R. J. St Leger, unpublished data). Pr1A was employed to produce an enhanced pathogenic strain of *M. a.* sf. *anisopliae* (St Leger *et al.*, 1996a). We identified seven additional subtilisins among the EST sequences of *M. a.* sf. *anisopliae* [Pr1C (AJ419628), Pr1D (AJ272861), Pr1G (AJ272863) Pr1H (AJ273173), Pr1I (AJ273780), Pr1J (AJ274368), Pr1G-K (AJ274144)], indicating the power of ESTs for mining the genome. These additional sequences were subsequently confirmed by cloning from genomic DNA (S. Bagga, S. Screen, F. M. Freimoser & R. J. St

Table 1. The 10 most abundant transcripts among the EST sequences for *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum*

M. a. sf. *anisopliae*	Abundance (%)	*M. a.* sf. *acridum*	Abundance (%)
Trypsin, Try 2	9·4	Subtilisin, Pr1A	7·4
Subtilisin, Pr1A	6·7	Chitinase, Chi 1	2·8
Trypsin, Try 1	2·5	Spore coat protein	1·6
Carboxypeptidase	2·1	Chitosanase, Csn 1	0·9
Clock-controlled protein	1·8	Chitinase, Chi 2	0·9
DNaseI	1·2	Chitinase, Chi 3	0·9
Control of DNA replication	1·2	Carboxypeptidase	0·9
rASP f4	1·1	Elongation factor	0·7
Cell wall protein	1·0	Unknown protein	0·5
Subtilisin, Pr1J	0·8	GlcN-6-phosphate isomerase	0·4

Leger, unpublished data). To date, the diversity of subtilisins is a unique feature of *M. anisopliae* and like other large gene families presumably arose by gene duplication followed by divergent evolution (Ohno, 1970). Although trypsins are among the most abundant transcripts in *M. a.* sf. *anisopliae*, like most of the subtilisins they were absent from the ESTs of *M. a.* sf. *acridum* (see Table B available as supplementary data at http://mic.sgmjournals.org). However, they could all be PCR-amplified from genomic DNA of *M. a.* sf. *acridum* (S. Bagga, S. Screen, F. M. Freimoser & R. J. St Leger, unpublished data). This suggests that differences in the nature and number of pathogenicity factors and in pathogenicity per se might be due to different regulation of the same set of genes rather than gain and loss of genes.

Subtilisins have been intensively studied in insect–fungus interactions (St Leger & Screen, 2001). Similarly, the aspartyl and metalloproteases from *M. anisopliae* have close homologues in genes expressed during infection processes by the plant pathogen *Glomerella cingulata* (Clark *et al.*, 1997) and the human pathogen *A. fumigatus* (Sirakova *et al.*, 1994), respectively. Trypsins are the most abundant proteases secreted by many plant-pathogenic ascomycetes (St Leger *et al.*, 1997). Most of the barriers and nutritional resources in the insect cuticle and in insect haemolymph are proteinaceous. The proteolytic array could function synergistically to achieve rapid physical ingress, nutrient solubilization and the disabling of antimicrobial peptides and thus may constitute quantitative factors that contribute to the overall virulence of the pathogen. With the broad array of different subtilisins found in a single strain, it is tempting to speculate that each type could have different biological properties and function.

Besides nutrition, host molecules and degradation products such as amino acids and peptides might also function as signalling molecules at the plasma membrane or following uptake (Paterson *et al.*, 1994). Included among various amino acid and peptide transporters expressed by *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum* are homologues to transporters involved in the uptake of host-derived nutrients by other pathogens. Thus ESTs AJ274326 and AJ272773 from *M. a.* sf. *anisopliae* are both highly similar ($E = 5 \times 10^{-96}$ and 2×10^{-62}, respectively) to the *inda*1 gene from *Trichoderma harzianum*, which encodes an amino acid transporter specifically induced during mycoparasitsm (Vasseur *et al.*, 1995). Several ESTs of *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum* had counterparts in amino acid and peptide transporters from the human pathogen *Candida albicans* while others had homologues in saprophytes such as *S. cerevisiae* and *N. crassa* (see the Tables available as supplementary data at http://mic.sgmjournals.org). Both *M. anisopliae* subspecies had homologues with ABC transporters, which in human pathogens are directly involved in contact-dependent secretion (including peptides), virulence and multiple drug resistance to antifungal compounds (Theiss *et al.*, 2002). In this context, as insects evolved before vertebrates and so did their pathogens, there is an interesting possibility that pathogens such as *M. anisopliae* include the progenitors of virulence factors in vertebrate pathogens.

It has been noted that fungi produce and secrete many enzymes that are toxic components of bacterial and animal venoms and are consequently potential virulence determinants (St Leger & Screen, 2000). Based on enzyme assays, proteases, glycosidic activities, esterases, phosphodiesterase, phospholipases, phosphatases and sulfatases were determined in the culture media of different fungi (St Leger & Screen, 2000). We found ESTs for most of these activities (see the Tables available as supplementary data at http://mic.sgmjournals.org). Thus, AJ274108 and BQ143695 are very similar to a phospholipase A ($E = 1 \times 10^{-45}$) and a phospholipase C ($E = 1 \times 10^{-51}$), respectively, and these activities are frequently reported to be bacterial virulence factors (Flieger *et al.*, 2001). A lysophospholipase that is most similar ($E = 4 \times 10^{-54}$) to an enzyme from *Penicillium chrysogenum* also has a homologue that contributes to development of legionnaires' disease in a number of ways, including protecting *Legionella pneumophila* from toxic products generated by other phospholipases (Flieger *et al.*, 2001). Tannases similar to those produced by some plant pathogens were tagged in both *M. a.* sf. *anisopliae* (AJ272848) and in *M. a.* sf. *acridum* (BQ142853). Along with peroxidases, laccases and monophenol oxygenase, these enzymes may be involved in mobilizing cuticular components rendered recalcitrant due to reactions with phenolic acids and a tanning process similar to that of plant cell walls (Neville, 1984). Both subspecies produced a variety of antioxidant proteins, including catalases and peroxidases that are involved in the pathogenicity of animals and plants by fungi (Wu *et al.*, 1997) and may also provide protection for invertebrate pathogens against active oxygen species generated as part of the host defence response (Iwanaga & Kawabata, 1998). Production of antimicrobials to defend against opportunistic competitors may also contribute to pathogenic strategies. EST BQ143138 from *M. a.* sf. *acridum* was homologous ($E = 8 \times 10^{-41}$) to a lysozyme from the fungus *Chalara* sp. (Felch *et al.*, 1975; Lyne *et al.*, 1990) and to EST sequences from *N. crassa* and *A. nidulans*.

The putative role of some of the ESTs encoding secreted enzymes is more ambiguous as obvious substrates are lacking in insect hosts. Some of these could facilitate saprophytic growth during the soil-dwelling component of the life-cycle of *M. a.* sf. *anisopliae*. Thus *M. a.* sf. *anisopliae* (AJ273623) had its closest counterpart in a β-glucosidase from *Phaeosphaeria avenaria* ($E = 3 \times 10^{-51}$) that hydrolyses plant saponins (Morrissey *et al.*, 2000) while ESTs BQ143361 and BQ143643 were similar to pectin-degrading polygalacturonases ($E = 8 \times 10^{-31}$ and 6×10^{-32}, respectively) (Scott-Craig *et al.*, 1998; Wubben *et al.*, 1999). Both subspecies showed several hits with transposase-like sequences (e.g., AJ273429, AJ274202, BQ143622) along with polyproteins (see the Tables available as supplementary

data at http://mic.sgmjournals.org), indicating that transfer events were occurring. Insertional mutagenesis events have obvious implications for strain stability that are of importance when considering the commercial development of a strain and the possibility of alterations in virulence and host range.

Other products besides enzymes are secreted by *M. anisopliae* and may be crucial for pathogenicity. Germ tubes secrete copious amounts of adhesive mucilage at the germ tube tip that also functions as an environment for secreted enzymes (St Leger, 1993). The composition of the mucilage is unknown, but EST AJ272837, with homology ($E = 2 \times 10^{-23}$) to a mucin-like glycoprotein that mediates invasion by *Cryptosporidium parvum* (Barnes *et al.*, 1998), suggests an avenue for investigation.

The products of some pathogenicity genes will be involved in the exchange of signals between the pathogen and its host, and activation of pathogenic mechanisms. As an example, *M. anisopliae* uses enzymes expressed at low levels to sense the nature of the polymeric nutrient present in the immediate environment (Screen *et al.*, 2002; St Leger *et al.*, 1986a). Likewise, the plant pathogen *Fusarium solani* penetrates plant cell walls using cutinase, produced in response to soluble monomers released by constitutive production of cutinase (Li *et al.*, 2002). Production by *M. anisopliae* of homologues to cutinase transcription factor 1 (AJ272967) and cutinase G-box binding protein (AJ274235) ($E = 3 \times 10^{-76}$ and 1×10^{-63}, respectively), previously only known to be involved in cutinase induction in *F. solani* (Li *et al.*, 2002), implies similarities between the regulatory circuitry of these pathogens with very different hosts. As there is no indication of the production of cutinases by *M. a.* sf. *anisopliae* (St Leger *et al.*, 1997), these elements may be involved in regulating expression of a range of secreted virulence factors in different fungi. Other *M. a.* sf. *anisopliae* genes similar to genes from plant-associated fungi include AJ272824 similar ($E = 7 \times 10^{-53}$) to MAS1 (AF264035) from *Magnaporthe grisea* produced during appressorium formation and AJ273567 similar ($E = 2 \times 10^{-56}$) to a gene from the ectomycorrhizal fungus *Laccaris bicolar* that is required for the initiation and maintenance of symbiosis (Kim *et al.*, 1998). These elements imply the existence of previously unsuspected components in pathogen–insect interactions and some shared feature(s) of fungal biology that are fundamental and possibly pre-adaptive in that they may make transitions to pathogenicity or transitions between very different hosts relatively simple. It is also possible that very similar genes may have evolved different roles as a result of selection in different genetic backgrounds in which case divergent selection acting on a few key traits may have played an important role in the evolution of fungi. In any event, these genes would provide a conceptual framework to guide investigations of specific host–pathogen interactions.

There were many EST sequences that had non-fungal sequences as their best BLAST hit. These included many antimicrobial molecules that may provide *M. anisopliae* with a selective advantage in defending limited resources within soil or the insect cadaver. EST AJ273066 from *M. a.* sf. *anisopliae* showed the most similarity with anti-fungal plant thaumatins, e.g., from *Arabidopsis thaliana* ($E = 10^{-25}$) (Selitrennikoff, 2001). However, while some fungi express thaumatin-like activity (Grenier *et al.*, 2000), the *M. a.* sf. *anisopliae* sequence shows weak homology ($E = 4 \times 10^{-6}$) to the only fungal sequence currently in the databases [from the plant pathogen *Glomerella cingulata* (AAL78508)]. Several other *M. anisopliae* sequences showed weaker similarity with plant antimicrobials. Thus, EST AJ273439 was most closely related to a phenylcoumaran benzylic ether reductase from *Pinus taeda* ($E = 2 \times 10^{-6}$) that produces antifungal toxins (Gang *et al.*, 1999).

Among the more challenging ESTs are the ones most similar to bacterial and animal sequences but with an extraordinarily patchy distribution. Thus AJ274050 only shows homologies to p67-phox in some mammals, particularly that of the dolphin *Tursiops truncatus* ($E = 7 \times 10^{-23}$). These produce superoxide ions in response to microbial infections (Bunger *et al.*, 2000). One EST clone of *M. a.* sf. *anisopliae* (AJ273180) was similar to a gene involved in the biosynthesis of phenazine ($E = 9 \times 10^{-13}$), a known pathogenicity factor in *Pseudomonas aeruginosa* (Mahajan *et al.*, 1999). Other *M. anisopliae* ESTs that were much more similar to bacterial sequences than to the closest fungal match included homologues to a catalase peroxidase (BQ143330, $E = 7 \times 10^{-57}$) and a sugar hydrolase (BQ143342, $E = 8 \times 10^{-20}$), both in *M. a.* sf. *acridum*, and several hypothetical, unknown and uncharacterized proteins. This suggests that some genes transcribed by *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum* could have been acquired via horizontal gene transfer. However, to strengthen or discard the hypothesis of horizontal transfer a much more detailed analysis of their phylogenetic relationship and distribution in other fungal lineages is required. Such an analysis recently made a case for horizontal transfer of chymotrypsin from a streptomycete to *M. a.* sf. *anisopliae* (Screen & St Leger, 2000).

Highly represented transcripts vary between *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum*

Clear differences were observed in the patterns of gene expression between *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum*. In general, genes necessary for the synthesis or degradation of cell walls and carbohydrate metabolism were more abundant in *M. a.* sf. *acridum* (Table 1, Fig. 1). Among the 10 most frequent transcripts in *M. a.* sf. *acridum* were three chitinases and a chitosanase (Table 1), presumably reflecting both its greater propensity to produce chitinases (St Leger *et al.*, 1996c) and induction by the additional chitin in its growth medium. Chitinases have been characterized in several insect-pathogenic fungi and are produced

during host cuticle penetration (St Leger *et al.*, 1993a, 1996c).

M. a. sf. *acridum* expressed more genes involved in stress response, detoxification and transmembrane transport (Fig. 1) than *M. a.* sf. *anisopliae*. In contrast *M. a.* sf. *anisopliae* expressed more genes that encoded enzymes involved in the synthesis of toxic metabolites and the control of the cell cycle and growth than *M. a.* sf. *acridum* (Fig. 1). In part, these may derive from differences in growth conditions as well as fungal genotype as we derived cDNA populations under different host-related conditions. However, the classes of ESTs that were found mirror differences in the ecology of *M. a.* sf. *anisopliae* and *M. a.* sf. *acridum*. Thus, the presence of several transcripts encoding enzymes involved in the synthesis of toxic metabolites in *M. a.* sf. *anisopliae* and the absence of counterparts in *M. a.* sf. *acridum* is also representative of the different strategies these two fungi use. *M. a.* sf. *acridum* invades all tissues of the host and the insect dies when it is filled with fungal biomass (Inglis *et al.*, 2001), consistent with an absence of toxins. In contrast, several *M. a.* sf. *anisopliae* ESTs were similar to peptide synthases, reductases and other enzymes that take part in the synthesis of fungal toxins such as aflatoxin, destruxins, trichothecene and enniatin (see Table B available as supplementary data at http://mic.sgmjournals.org). This is in agreement with the observation that *M. a.* sf. *anisopliae* rapidly kills its host after infection through the action of toxins and consequently colonizes the insect host by saprobic growth (Samuels *et al.*, 1989). There were more ESTs in *M. a.* sf. *anisopliae* with counterparts involved in light perception and circadian signalling (Fig. 1), providing the first evidence that *M. anisopliae* may react to light. Unlike *M. a.* sf. *anisopliae*, *M. a.* sf. *acridum* sporulates within insects as an adaptation to desert living (Inglis *et al.*, 2001), which could be related to differences in light perception.

We have presented and discussed here only an initial analysis of the EST dataset and further characterized selected examples with emphasis on secreted products with putative roles in host invasion. The amount of redundancy present in the EST dataset is relatively low. It is therefore likely that the generation of more sequence data will identify more novel pathogenicity-related genes. However, this comparative EST project has already provided a global perspective to the biology and biosynthetic capacity of the ubiquitous insect pathogen *M. anisopliae*, and a new foundation for analysis of gene function with clones for more than 2000 different genes available. Rational strain improvement has been brought closer and a comprehensive expression analysis will determine which subset of the identified genes is expressed during host cuticle invasion. The study also confirmed and revealed differences and similarities between two closely related but ecologically distinct fungal pathogens, and has allowed an initial comparison with plant pathogens and non-pathogens. Many of the genes identified in *M. anisopliae* suggest similarities of fungal infections, regardless of host. In general, genes that contribute to ecological diversification and the nature of the forces acting during this process are poorly known, partly because genes involved in ecological attributes are hard to identify (Duda & Palumbi, 1999). The knowledge gained through this study on secreted proteins will help remedy this deficiency and the collection of ESTs will also provide a resource that will enable us to address questions about the evolution, regulation and networking of pathogenicity determinants in the two subspecies of *M. anisopliae*.

ACKNOWLEDGEMENTS

This work was supported by the University of Maryland, USDA (grant 2001-35302-10141) and the NSF (grant MCB-0091196). F. M. F. acknowledges support from the Swiss National Science Foundation. We would like to thank A. Wilson for computer support.

REFERENCES

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **17**, 3389–3402.

Anderson, I. & Brass, A. (1998). Searching DNA databases for similarities to DNA sequences: when is a match significant? *Bioinformatics* **14**, 349–356.

Barnes, D. A., Monnin, A., Huang, J. X., Gousset, L., Wu, J., Gut, J., Doyle, P., Dubremetz, J. F., Ward, H. & Petersen, C. (1998). A novel multi-domain mucin-like glycoprotein of *Cryptosporidium parvum* mediates invasion. *Mol Biochem Parasitol* **96**, 93–110.

Bunger, P. L., Swain, S. D., Clements, M. K., Siemsen, D. W., Davis, A. R., Gauss, K. A. & Quinn, M. T. (2000). Cloning and expression of bovine p47-phox and p67-phox: comparison with the human and murine homologs. *J Leukoc Biol* **67**, 63–72.

Clark, S. J., Templeton, M. D. & Sullivan, P. A. (1997). A secreted aspartic proteinase from *Glomerella cingulata*: Purification of the enzyme and molecular cloning of the cDNA. *Microbiology* **143**, 1395–1403.

Crameri, R. & Blaser, K. (1996). Cloning *Aspergillus fumigatus* allergens by the pJuFo filamentous phage display system. *Int Arch Allergy Immunol* **110**, 41–45.

Duda, T. F. & Palumbi, S. R. (1999). Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. *Proc Natl Acad Sci U S A* **96**, 6820–6823.

Felch, J. W., Inagami, T. & Hash, J. H. (1975). The *N,O*-diacetylmuramidase of *Chalaropsis* species. V. The complete amino acid sequence. *J Biol Chem* **250**, 3713–3720.

Flieger, A., Gong, S., Faigle, M., Stevanovic, S., Cianciotto, N. P. & Neumeister, B. (2001). Novel lysophospholipase A secreted by *Legionella pneumophila*. *J Bacteriol* **183**, 2121–2124.

Gang, D. R., Kasahara, H., Xia, Z.-Q., Vander Mijnsbrugge, K., Bauw, G., Boerjan, W., Van Montagu, M., Davin, L. B. & Lewis, N. G. (1999). Evolution of plant defense mechanisms. Relationship of phenylcoumaran benzylic ether reductases to pinoresinol-lariciresinol and isoflavone reductases. *J Biol Chem* **274**, 7516–7527.

Grenier, J., Potvin, C. & Asselin, A. (2000). Some fungi express beta-1,3-glutanases similar to thaumatin-like proteins. *Mycologia* **92**, 841–848.

Huang, X. & Madan, A. (1999). CAP3: a DNA sequence assembly program. *Genome Research* **9**, 868–877.

Inglis, G. D., Goettel, M. S., Butt, T. M. & Strasser, H. (2001). Use of hyphomycetous fungi for managing insect pests. In *Fungi as Biocontrol Agents*, pp. 23–69. Edited by T. M. Butt, C. Jackson & N. Magan. Wallingford: CAB International.

Iwanaga, S. & Kawabata, S. (1998). Evolution and phylogeny of defense molecules associated with innate immunity in horseshoe crab. *Front Biosci* **3**, 973–984.

Joshi, L. & St Leger, R. J. (1999). Cloning, expression, and substrate specificity of MeCPA, a zinc carboxypeptidase that is secreted into infected tissues by the fungal entomopathogen *Metarhizium anisopliae*. *J Biol Chem* **274**, 9803–9811.

Joshi, L., St Leger, R. J. & Roberts, D. W. (1997). Isolation of a cDNA encoding a novel subtilisin-like protease (Pr1B) from the entomopathogenic fungus, *Metarhizium anisopliae* using differential display-RT-PCR. *Gene* **197**, 1–8.

Kim, S.-J., Zheng, J., Hiremath, S. T. & Podila, G. K. (1998). Cloning and characterization of a symbiosis-related gene from an ectomycorrhizal fungus *Laccaria bicolor*. *Gene* **222**, 203–212.

Li, D., Sirakova, T., Rogers, L., Ettinger, W. F. & Kolattukudy, P. E. (2002). Regulation of constitutively expressed and induced cutinase genes by different zinc finger transcription factors in *Fusarium solani f.* sp. *pisi* (*Nectria haematococca*). *J Biol Chem* **277**, 7905–7912.

Lyne, J. E., Carter, D. C., Xiao-min, H., Stubbs, G. & Hash, J. H. (1990). Preliminary crystallographic examination of a novel fungal lysozyme from *Chalaropsis*. *J Biol Chem* **265**, 6928–6930.

Mahajan, M. S., Tan, M. W., Rahme, L. G. & Ausubel, F. M. (1999). Molecular mechanisms of bacterial virulence elucidated using a *Pseudomonas aeruginosa–Caenorhabditis elegans* pathogenesis model. *Cell* **96**, 47–56.

Morrissey, J. P., Wubben, J. P. & Osbourn, A. E. (2000). *Stagnospora avenae* secretes multiple enzymes that hydrolyze oat leaf saponins. *Mol Plant–Microbe Interact* **13**, 1041–1052.

Nelson, M. A., Kang, S. C., Braun, E. L. & 23 other authors (1997). Expressed sequences from conidial, mycelial, and sexual stages of *Neurospora crassa*. *Fungal Genet Biol* **21**, 348–363.

Neville, A. C. (1984). Cuticle: organisation. In *Biology of the Integument. 1. Invertebrates*, pp. 611–625. Edited by J. Breiter-Hahn, A. G. Matoltsy & K. S. Richards. Berlin: Springer.

Ohno, S. (1970). *Evolution by Gene Duplication*. Berlin, New York: Springer.

Paterson, I. C., Charnley, A. K., Cooper, R. M. & Clarkson, J. M. (1994). Partial characterization of specific inducers of a cuticle-degrading protease from the insect pathogenic fungus *Metarhizium anisopliae*. *Microbiology* **140**, 3153–3159.

Reddy, P. V., Lam, C. K. & Belanger, F. C. (1996). Mutualistic fungal endophytes express a proteinase that is homologous to proteases suspected to be important in fungal pathogenicity. *Plant Physiol* **111**, 1209–1218.

Rippon, J. W. (1988). *Medical Mycology: the Pathogenic Fungi and the Pathogenic Actinomycetes*, 3rd edn. Philadelphia: Sanders.

Roberts, D. W. & Humber, R. A. (1981). Entomogenous fungi. In *Biology of Conidial Fungi*, pp. 201–236. New York: Academic Press.

Samuels, K. D. Z., Pinnock, D. E. & Allsopp, P. G. (1989). The potential of *Metarhizium anisopliae* (Metschnikoff) Sorokin (Deuteromycotina, Hyphomycetes) as a biological control-agent of *Inopus rubriceps* (Macquart) (Diptera, Stratiomyidae). *J Aust Entomol Soc* **28**, 69–74.

Scott-Craig, J. S., Cheng, Y. Q., Cervone, F., De Lorenzo, G., Pitkin, J. W. & Walton, J. D. (1998). Targeted mutants of *Cochliobolus carbonum* lacking the two major extracellular polygalacturonases. *Appl Environ Microbiol* **64**, 1497–1503.

Screen, S. E. & St Leger, R. J. (2000). Cloning, expression, and substrate specificity of a fungal chymotrypsin. Evidence for lateral gene transfer from an actinomycete bacterium. *J Biol Chem* **275**, 6689–6694.

Screen, S. E., Hu, G. & St Leger, R. J. (2002). Transformants of *Metarhizium anisopliae* sf. *anisopliae* overexpressing chitinase from *Metarhizium anisopliae* sf. *acridum* show early induction of native chitinase but are not altered in pathogenicity to *Manduca sexta*. *J Invertebr Pathol* **78**, 260–266.

Selitrennikoff, C. P. (2001). Antifungal proteins. *Appl Environ Microbiol* **67**, 2883–2894.

Sirakova, T. D., Markaryan, A. & Kolattukudy, P. E. (1994). Molecular cloning and sequencing of the cDNA and gene for a novel elastinolytic metalloproteinase from *Aspergillus fumigatus* and its expression in *Escherichia coli*. *Infect Immun* **62**, 4208–4218.

St Leger, R. (1993). Biology and mechanisms of insect cuticle invasion by deuteromycete fungal pathogens. In *Parasites and Pathogens of Insects*, pp. 211–229. Edited by N. C. Beckage, S. N. Thompson & B. A. Federici. New York: Academic Press.

St Leger, R. & Bidochka, M. J. (1996). Insect–fungal interactions. In *New Directions in Invertebrate Immunology*, pp. 443–479. Edited by K. Söderhäll, S. Iwanaga & G. Vasta. New Haven, NJ: SOS Publications.

St Leger, R. J. & Screen, S. E. (2000). *In vitro* utilization of mucin, lung polymers, plant cell walls and insect cuticle by *Aspergillus fumigatus*, *Metarhizium anisopliae* and *Haematonectria haematococca*. *Mycol Res* **104**, 463–471.

St Leger, R. & Screen, S. (2001). Prospects for strain improvement of fungal pathogens of insects and weeds. In *Fungal Biocontrol Agents: Progress, Problems and Potential*, pp. 219–238. Edited by T. M. Butt, C. Jackson & N. Morgan. Wallingford: CAB International.

St Leger, R. J., Charnley, A. K. & Cooper, R. M. (1986a). Cuticle-degrading enzymes of entomopathogenic fungi: mechanisms of interaction between pathogen enzymes and insect cuticle. *J Invertebr Pathol* **47**, 295–302.

St Leger, R. J., Charnley, A. K. & Cooper, R. M. (1986b). Cuticle-degrading enzymes of entomopathogenic fungi: synthesis in culture on cuticle. *J. Invertebr Pathol* **48**, 85–95.

St Leger, R. J., Charnley, A. K. & Cooper, R. M. (1987). Characterization of cuticle-degrading proteases produced by the entomopathogen *Metarhizium anisopliae*. *Arch Biochem Biophys* **253**, 221–232.

St Leger, R. J., Butt, T. M., Staples, R. C. & Roberts, D. W. (1989). Synthesis of proteins including a cuticle-degrading protease during differentiation of the entomopathogenic fungus *Metarhizium anisopliae*. *Exp Mycol* **13**, 253–262.

St Leger, R. J., May, B., Allee, L. L., Frank, D. C., Staples, R. C. & Roberts, D. W. (1992). Genetic differences in allozymes and in formation of infection structures among isolates of the entomopathogenic fungus *Metarhizium anisopliae*. *J Invertebr Pathol* **60**, 89–101.

St Leger, R. J., Staples, R. C. & Roberts, D. W. (1993a). Entomopathogenic isolates of *Metarhizium anisopliae*, *Beauveria bassiana* and *Aspergillus flavus* produce multiple extracellular chitinase isozymes. *J lnvertebr Pathol* **61**, 81–84.

St Leger, R. J., Cooper, R. M. & Charnley, A. K. (1993b). Analysis of aminopeptidase and dipeptidylpeptidase from the entomopathogenic fungus *Metarhizium anisopliae*. *J Gen Microbiol* **139**, 237–243.

St Leger, R. J., Bidochka, M. J. & Roberts, D. W. (1994). Isoforms of the cuticle-degrading Pr1 proteinase and production of a metalloproteinase by *Metarhizium anisopliae*. *Arch Biochem Biophys* **313**, 1–7.

St Leger, R., Joshi, L., Bidochka, M. J. & Roberts, D. W. (1996a). Construction of an improved mycoinsecticide overexpressing a toxic protease. *Proc Natl Acad Sci U S A* **93**, 6349–6354.

St Leger, R. J., Joshi, L., Bidochka, M. J., Rizzo, N. W. & Roberts, D. W. (1996b). Biochemical characterization and ultrastructural localization of two extracellular trypsins produced by *Metarhizium anisopliae* in infected insect cuticles. *Appl Environ Microbiol* **62**, 1257–1264.

St Leger, R. J., Joshi, L., Bidochka, M. J., Rizzo, N. W. & Roberts, D. W. (1996c). Characterization and ultrastructural localization of chitinases from *Metarhizium anisopliae*, *M. flavoviride*, and *Beauveria bassiana* during fungal invasion of host (*Manduca sexta*) cuticle. *Appl Environ Microbiol* **62**, 907–912.

St Leger, R. J., Joshi, L. & Roberts, D. W. (1997). Adaptation of proteases and carbohydrases of saprophytic, phytopathogenic and entomopathogenic fungi to the requirements of their ecological niches. *Microbiology* **143**, 1983–1992.

Theiss, S., Kretschmar, M., Nichterlein, T., Hof, H., Agabian, N., Hacker, J. & Kohler, G. A. (2002). Functional analysis of a vacuolar ABC transporter in wild-type *Candida albicans* reveals its involvement in virulence. *Mol Microbiol* **43**, 571–584.

Vasseur, V., Van Montagu, M. & Goldman, G. H. (1995). *Trichoderma harzianum* genes induced during growth on *Rhizoctonia solani* cell walls. *Microbiology* **141**, 767–774.

Wu, G. S., Shortt, B. J., Lawrence, E. B., Leon, J., Fitzsimmons, K. C., Levine, E. B., Raskin, I. & Shah, D. M. (1997). Activation of host defense mechanisms by elevated production of H_2O_2 in transgenic plants. *Plant Physiol* **115**, 427–435.

Wubben, J. P., Mulder, W., ten Have, A., van Kan, J. A. & Visser, J. (1999). Cloning and partial characterization of endopolygalacturonase genes from *Botrytis cinerea*. *Appl Environ Microbiol* **65**, 1596–1602.