ABSTRACT

Title of Dissertation: USING THE FUNGUS METARHIZIUM ANISOPLIAE AS A MODEL SYSTEM TO STUDY THE ROLE OF GENE DUPLICATION, DIVERGENCE AND EXPRESSION IN ADAPTING TO PATHOGENICITY.

Gang Hu , Doctor of Philosophy 2005

Directed By: Professor Raymond J. St. Leger
Department of Entomology

Fungal pathogens have been recorded for virtually all groups of multicellular organisms. They are the major cause of disease in insects in nature, sometimes causing devastating epizootics. Entomopathogenic fungi are applied in the field as biological control agents world-wide. However, there remain many gaps in our knowledge concerning pathogenicity of fungi towards insects (St. Leger and Screen 2001).

This study focuses on the entomopathogenic fungus *Metarhizium anisopliae*. In particular, it focuses on the evolutionary significance and biocontrol potential of gene duplication and loss events among its plethora of serine proteases. The serine proteases produced by *M. anisopliae* and other pathogenic fungi currently provide the best-understood model of fungal determinants of pathogenicity (St. Leger and Screen 2001). The importance of subtilase Pr1A during the progress of infection by *M. anisopliae* was first identified through its high concentration at the site of

penetration, and its considerable ability to degrade the cuticular integument (St. Leger *et al* 1987, 1989). Here a phylogenomic approach is adopted with fungi of very different virulence and habitat to survey and characterize their serine proteinases with the goal of providing a framework of information on these important enzymes, as well as improving understanding of general processes in fungal gene family evolution.

Serine proteases became candidate for gene manipulation due to their critical role in the infection process. A strain of *M. anisopliae* was genetically modified so it overexpressed Pr1A and kills insects faster than the wild type does in laboratory tests (St. Leger *et al*. 1996). This technology has potential for pest control (St. Leger 2001), but there is an uncertainty about the efficacy, survivability, and environmental risk posed by any introduced or engineered fungus because of our lack of knowledge about the fate of fungal genotypes at the population and ecosystem levels (Bidochka 2001, Hajek *et al*. 1997). So a planned release (approved by EPA) was conducted in a field of cabbage plants. The release established the technology required to monitor the fate of genetically enhanced *M. anisopliae*. This technology was used to determine the potential of engineered strains to establish and disperse over 1-year test period.

USING THE FUNGUS METARHIZIUM ANISOPLIAE AS A MODEL SYSTEM TO STUDY THE ROLE OF GENE DUPLICATION, DIVERGENCE AND EXPRESSION IN ADAPTING TO PATHOGENICITY.

By

Gang Hu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2005

Advisory Committee:
Professor Raymond J. St. Leger, Chair
Associate Professor Earlene Armstrong
Professor Michael Co Ma
Assistant Professor Louisa Wu
Professor Robert Yuan

Dedication

To my parents, Zhaohui Hu and Huaiwen Yang and my wife Bo Liu. It is they who took me on this route; And now I hope, they'll enjoy the fruit.

Acknowledgements

I would like to thank all the people who helped me with my research. I would like to thank the members of my advisory committee: Dr. Earlene Armstrong, Dr. Michael Co Ma, Dr. Luisa Wu and Dr. Robert Yuan. I would like to express my deepest appreciation to my supervisor Prof. Raymond J. St. Leger for his support, guidance, patience, encouragement and friendship.

And I want to thank the faculties in department of Entomology for their encouragement during the course of my study.

Table of Contents

List of Tables

List of Figures

.

Chapter 1: Introduction

Insect pathologists initially focused their studies on protecting beneficial insects from fungal diseases. Now the focus has moved on. There is considerable interest in using fungal pathogens to control insect pests. However, while ninety fungal genera, and more than seven hundred species, are pathogenic to insects (Roberts & Humber 1981) only a few fungal genera are currently being explored for insect pest control, namely species of *Entomophaga, Neozygites, Entomophthora, Erynia, Aschersonia, Verticillium, Nomuraea, Hirsutella, Metarhizium, Beauveria* and *Paecilomyces.*

Literature review

The biology of fungal pathogens

Since the pathogenic fungi are found in many divergent fungal groups their ability to infect insects must have evolved independently several times. Fungi may have emerged as endophytes of plants (Lewis 1987). The insect pathogenic fungi may have evolved from either these endophytes or from the plant pathogenic fungi (themselves descendents of endophytes) via adaptations in the extra cellular hydrolytic enzymes to accommodate hydrolysis of proteinaceous insect cuticles (St. Leger & Bidochka 1996). Unlike other microbial insect pathogens, such as bacteria, virus and protozoa, which infect the host via ingestion, entomopathogenic fungi, invade the host by penetration of the outer integument. This enables parasitization of non-feeding, pupal and egg stages as well as the many sap and blood-sucking insects (Chamley 1989a).

There are some similarities in the overall structures of the insect cuticle and plant integuments in so far as both are composite structures containing a fibrous material (chitin or cellulose) embedded in the matrix material (protein or pectic substances and hemicelluloses) (Chamley 1984, Carpita & Gibeaut 1993). This suggests that the overall mechanism of fungal pathogenesis against insects and plants may be similar. Fungal pathogens of both groups have evolved highly sophisticated regulatory mechanisms to allow them to detect changes in their environment and respond accordingly. St. Leger *et al.* (1997) have examined the production of secreted enzymes in the cultures of representative saprophytic, entomopathogenic and phytopathogenic fungi. According to their study, saprophytes produced the broadest range of enzymatic activities compared to pathogens that produced a subset of enzymes reflecting their ecological niches. However, for most fungal-host interactions, the nature of its genes and their products, from both hosts and pathogens that limit the host range and the host's resistance and susceptibility are poorly understood.

Metarhizium anisopliae, a pathogen with worldwide distribution has been isolated from more than 200 insect species across 7 orders (Table 1.). *Metarhizium sp*. shows great potential as a biocontrol agent, and is under development in several countries as a control agent against many species (Table 2). Strains of *Metarhizium* differ in their host range, necessitating selection of the most virulent strain against a target insect (Zimmerman 1993).

Table 1. Insect orders from which *M. anisopliae* has been isolated (Zimmerman 1993).

Family	Examples	Number of *M. anisopliae* strains isolated
Orthoptera	crickets, grasshoppers	11
Derinaptera	earwigs	1
Hemiptera	bugs, leafhoppers	21
Lepidoptera	butterflies, moths	17
Diptera	flies	4
Hymenoptera	bees, wasps, sawflies	6
Coleoptera	beetles	134

Table 2 Development of *M. anisopliae* as a biocontrol agent, insect species and locations (Zimmerman 1993).

Family	Target organism	Country
Orthoptera	locusts, grasshoppers, *Schistocerca gregaria*, *Locusta migratoria R & F*. Cockroaches, *Periplaneta americans L.*	USA
Isoptera	Termites	Australia
Homoptera	Spittlebugs, *Aenolamia spp.*	Australia
	Nila arvata lugens,	Philippines
	Mohanarva posticata, Zilia spp.	Brazil
Coleoptera	Chrysomelids, *Brontispa longissima Gest*.	Taiwan
	Curculionids *Otiorhynchus suicatus F.*	Europe
	Scarabaeids, *Adoryphorus couloni*	Australia
	Antirogus parvulus, Aphodius tasmaniae	
	Oryctes rhinoceros L	South Pacific, Philippines.

Mode of action of entomopathogens

Most of the information concerning pathogenicity of fungi towards insects has been derived from studies of the ascomycete fungi *M. anisopliae* and *Beauveria bassiana*.

For most entomopathogens the infectious propagule is the conidium/spore, which attaches to the host surface, germinates, and the germ tube penetrates the exoskeleton into the insect haemolymph. A wide array of spores and spore structures is observed among the pathogenic fungi and reflects their adaptation to the diversity of both the target hosts and their respective habitats.

The attachment of the fungal propagule to the host cuticle is mediated by the chemical components present on the outer layer of the spore wall and the epicuticle. Following attachment, the next critical stage in pathogenesis is germination and germ tube formation, as conidia on the surface of a host may suffer desiccation, antibiosis from saprophytic microbes or removal with the old cuticle if moulting occurs.

Conidia of *M. anisopliae* germinate in response to a range of exogenous carbon and nitrogen sources (St. Leger 1991). There is variation in germination triggers between different strains, which may be related to host species (St. Leger *et al*. 1994a). *Metarhizium anisopliae var. anisopliae* and *var. majus* isolated from scarabeids germinate at high frequency only in the presence of a crude protein/chitin product. Those isolated from other Coleoptera germinate and form appressoria in yeast extract (0.0125%) (St. Leger *et al*. 1994a). Isolates from hemiptera germinate and form appressoria in media containing glucose as the sole carbon source. The

order hemiptera includes aphids and other sap-sucking bugs. Aphids often become coated in the sugary sap, or honeydew released by their feeding, leading to high concentrations of carbon sources on the cuticle. A successful pathogen on these insects would therefore have to be capable of forming appressoria in the presence of high levels of available carbon. Evidently, nutrient levels are one of the environmental cues enabling the pathogen to 'recognize' if it has landed on an appropriate host cuticle. The level of nutrients on these comparatively thin skinned insects such as homopteran and lepidopteran larvae are likely to be higher than those on coleopterans with hard cuticles, thus leading to different responses (St. Leger *et al*. 1994a).

Differentiation of appressoria by *M. anisopliae* germ tubes can be induced in vitro when conidia germinate on a flat hydrophobic surface in a sparse nutritional environment (St. Leger. *et al*. 1989a). Germination and appressorial formation are not necessarily linked. Conidia applied to insects in inocula supplemented with nutrients (1% alanine or 1% N-acetylglucosaimine) failed to penetrate cuticle. Extensive growth of non-differentiated hyphae resulted in the formation of a thick mycelial mat over the cuticle surface. In the absence of extraneous nutrients, germlings showed directed growth into and through the cuticle (St. Leger *et al*. 1989a).
Penetration of the cuticle is thought to occur by a combination of enzymatic degradation and physical pressure. Evidence for enzymatic degradation includes the disappearance of the wax layer beneath appressoria of *M. anisopliae* on wireworm cuticle (Zacharuk 1970), and the presence of cuticular holes around germ tubes of *B. bassiana* during penetration of larvae of *Heliothis zea* (Pekrul & Grula, 1979).

Physical penetration has been shown during host invasion by some entomopthoralean pathogens (Brobyn & Wilding, 1983, Butt 1987). Enzymatic degradation of host cuticle may result from either extracellular or cell wall bound enzymes. Extracellular enzymes corresponding to the main chemical constituents of insect cuticle, protein, chitin and lipids, had been detected prior to 1986. However, there was little information about types, modes of action, regulation, cellular localization, sequence and levels of production (St. Leger *et al*. 1986a). Pathogenic isolates of *M. anisopliae, B. bassiana*, and *V. lecanii* grown in buffered and unbuffered liquid cultures containing 1% ground cuticle as the sole carbon source produced a range of extra-cellular cuticle-degrading enzymes (St. Leger *et al.* 1986a). Major enzyme activities included endoprotease, chitinase and N-acetylglucosaminidase (NAGase). Treatment of cuticle with these enzymes showed that pre-treatment with protease enhanced chitinase activity ca. 3.5 fold, implying the shielding of cuticular chitin by protein (St. Leger *et al.* 1986a). A similar co-operative effect was seen using commercial enzyme preparations to digest cuticle of the corn earworm, *Heliothis zea* (Smith *et al.* 1981). The sequence of enzymes produced in cuticle cultures by each species was similar. Esterase activity and enzymes of the proteolytic complex (protease, aminopeptidase and carboxypeptidase) appeared first (<24 hr) and increased rapidly after 28 hr. Chitinase was produced at low basal levels for 3 to 5 days, after which it increased rapidly. Lipase was not detected until day 5 for any fungus tested (St. Leger *et al*. 1986b). The early production of high levels of protease was considered an indication of their potentially important role in cuticle penetration and thus pathogenicity.

Role of proteolytic enzymes in fungal infection process

Proteases, especially serine proteases, have been studied extensively for their biochemical properties, mode of action and physiological functions. They are involved in many cellular processes such as protein turnover, cell nutrition, spore germination and enzyme secretion (Phadtare *et al.* 1997, St. Leger and Screen 2001). There is growing evidence that these enzymes are important pathogenicity determinants (St. Legers *et al*. 1996, St. Leger *et al*.1997).

St. Leger (1987a) detected two subtilisins and three trypsin-like proteases, and subsequently purified the major form of the subtilisin (Prl) and one form of the trypsin-like enzyme (Pr2). The cDNA encoding Pr1 was cloned by hybridization to a degenerate oligonucleotide designed against the N-terminal sequence (St. Leger *et al.* 1992b). The Pr1 gene encodes a 388 amino acid peptide, containing an 18 aa signal peptide, to allow translocation across the membrane, and a propeptide, to prevent proteolytic activity within the fungal cell. The mature protein shows similarity to members of the subtilisin family of serine proteases, with 65% identity to *Tritirachium album* proteinase K (St. Leger *et al*. 1992b). Prl production has been linked to the formation of appressoria and hyphae penetration. Production of Prl has been linked to infection structures by histochemical localization using gold-labeled antibodies (Goettel *et al.* 1989, St. Leger 1993). Prl could be detected at low levels in dormant or germinating conidia, with higher levels detected after differentiation of appressoria and during formation of the penetration peg and hyphal bodies. Initially Prl was found localized to the hyphal cell wall and in the immediate vicinity of the fungal structures, but as infection advanced (>40 hr) the enzyme diffused throughout

the cuticle coinciding with breakdown of cuticle lamellar structure and digestion of the protein moieties (St. Leger 1993).

Prl production has also been linked to nutrient deprivation. On being transferred from nutrient rich to nutrition deficient /starvation media (no carbon or nitrogen sources) mycelia synthesized Prl and Pr2 in less than 2hr (St. Leger *et al.* 1988b). Synthesis of active enzyme was substantially reduced when inhibitors of transcription (actinomycin D and 8-azoguanine) or translation (cyclohexamide and puromycin) were added. This indicated that there was de novo synthesis of mRNA involved in protease regulation. The appearance of active Prl during starvation has been further linked to de novo synthesis of mRNA by studies on translatable mRNA species produced by *M. anisopliae* mycelia during starvation (St. Leger *et al*. 1991b). Prl transcript levels have also been assayed in infection structures. Ribonucleic acid (RNA) isolated from conidia, non-differentiating germlings and differentiating germlings was subjected to Northern analysis against a Prl probe. Only faint hybridization was detectable in conidia and non-differentiated cells, in comparison to which differentiated cells showed an approximately 35-fold increase (by densitometry) in Prl mRNA levels (St. Leger *et al.* 1992b). Inhibition of Prl did not affect spore viability or prevent growth and formation of appressoria on the cuticle surface, suggesting that inhibition reduced infection by limiting fungal penetration of the cuticle (St. Leger *et al.* 1988a).

Metarhizium has been transformed to benomyl resistance using a P-tubulin gene from *Neurospora crassa* (Beniier *et al.* 1989). Transformation has been successfully used to examine the role of Prl in the pathogenic processes of *M.*

anisopliae. Plasmids carrying the *A. nidulans BenA3* gene in the coding region of the *pr1* gene were constructed and transformed into *M. anisopliae* (St. Leger, 1994d). The novel strains lacking Prl showed a significant delay in pathogenicity, which only commenced with the production of other proteases, including a metalloprotease with specificities similar to that of Prl. Similar studies have been performed with the subtilisin *(sep)* gene in *Aspergillus flavus,* a fungus causing invasive lung infections in immuno-compromised patient. A knockout mutation of the *sep* gene did not affect the pathogenicity of the fungus but a compensatory up-regulation of a metalloproteinase gene *(mep* 20) was detected. Elevated levels of the metalloprotease were also detected in the culture when wild type, *A. flavus* was treated with serine protease inhibitor (Ramesh & Kolattukudy, 1996). These studies indicate that parasitic fungi may have the ability to compensate for loss of some proteolytic activities by expressing other hydrolytic enzymes. From an evolutionary perspective this might be a useful strategy for the pathogen since several hosts, including plants and insects are known to produce serine protease inhibitors (Peanasky *et al.* 1984).

Considerably less is known about the other major protease produced by *M. anisopliae*, the trypsin-like Pr2 (St. Leger *et al*. 1987a). Like Pr1, Pr2 is stable at room temperature over a wide pH range. The trypsin-like nature of Pr2 is shown by its inhibition by the specific trypsin reagents, Leupeptin and Tos-Lys-CH2CI (St. Leger *et al.* 1987a). The production of Pr2 by infection structures was demonstrated by immunogold stain (St. Leger e*t al.* 1996) although activity against cuticle was approximately 1/4 of that found for Prl (St. Leger *et al.* 1987b). Regulation of production of Pr2 has been investigated. Paterson (1993) found that three insoluble

protein sources (insect cuticle, elastin and collagen) and two soluble proteins (BSA and gelatin) induced Pr2. The exact role of Pr2 in pathogenesis is unclear and has not been investigated. However, during synthesis on cuticle, Pr2 activity can be detected prior to Prl (St. Leger *et al.* 1988a) indicating a possible role in either the scavenging of nutrients or the release of peptides, which can induce production of Prl (Paterson *et al.* 1994a, b).

Serine proteases with essential serine residues, acid iso-electric points and substrate specificities similar to that of *M. anisopliae* (ARSEF 2575(MEI)) Pr2 activities have been detected from cultures of other well-known entomopathogenic fungi including *B. bassiana*, *Verticillium lecanii*, *Nomuraea rileyi* and *Aschersonia aleyroidis* (St. Leger *et al.* 1987c, 1994c). This widespread production of Pr2-like enzymes also indicates some role in pathogenicity

Gene duplication and evolution

Subtilisins have been isolated form a wide range of organisms, including archaebacteria, eubacteria, fungi and higher eukaryotes (Markland and Smith 1971; Siezen *et al.* 1991). Siezen *et al*. (1991) divided these enzymes into class I and class II; based on characteristic amino acid sequence patterns. All fungal subtilisins characterized to date were reported to be homologous with proteinase K from *T. album* and were consequently grouped in class II (Savita *et al.* 2004). Class II subtilisins are usually extracellular and have a nutritional role (Grnkel and GAssen 1989; Monod *et al*. 1991; Burton *et al.* 1993). The bacterial subtilisins are grouped as class I and may be exocellular or endocellular.

Gene duplication and subsequent functional divergence of duplicate genes is an important mechanism for evolution of novel genes (Nei 1969, Ohno 1970). It is still debatable whether functional divergence in most organisms occurs by positive Darwinian selection that accelerates the fixation of advantageous mutations (largely nonsynonymous or amino-acid altering mutations) (Goodman *et al*. 1975) or largely by random fixation of neutral mutations (Kimura 1983). The neutralist perspective on molecular evolution maintains that the vast majority of mutations affecting gene function are neutral or deleterious. If most mutations in coding regions are neutral or deleterious, the neutral theory of molecular evolution (Kimura 1993) predicts that the number of amino-acid substitutions in a protein should be roughly a linear function of time. After a gene duplication event where both genes are retained, it predicts that original and duplicate genes diverge at clock-like rates. This prediction has been tested for coding sequences, but can be applied to the gene expression pattern (Wagner 2000). Duplicate gene pairs with more highly diverged sequences should show more highly diverged expression patterns. The greater the sequence similarity between two duplicated genes, the greater should be the similarity in the expression pattern between these genes. This can be tested in two ways: 1) follow the divergence in sequence and expression pattern of one gene pair in multiple related organisms, or 2) analyze a large sample of duplicated gene pairs with varying degrees of divergence within one organism (Wagner 2000). Gene duplication provides the material for the generation of new genes and subsequently new function. How the members of the duplication gene pairs acquire new functions are still largely unknown. The classical view is that selection is relaxed, because if one gene gains a mutation that alters

function, the remaining copy serves as a backup to retain the original function (Ohta 1987). The divergent copy is then free to acquire random substitutions, and by chance these substitutions may result in new tissue specificity or function. It is probable that the vast majority of gene duplicates are mutated out of function within a few million years, with the few survivors experiencing strong purifying selection. Although duplications rarely evolve new functions, the stochastic silencing of such genes may play a significant role in the passive origin of new species. But it is still not clear how often gene duplications arise and how frequently they evolve new functions.

The functional divergence of gene sequences can occur by positive Darwinian selection that accelerate the fixation of advantageous mutations or by random genetic drift of neutral mutations, which later induce a change in gene function when the environment or genetic background is altered (Zhang *et al*. 1998). In the later case the rate of non synonymous nucleotide substitution (dN) may also be enhanced because of relaxed functional constraints of redundant genes after gene duplication but the rate will never exceed that of synonymous (silent) substitutions (dS). Therefore a widely accepted method of testing the neutral and selection theories of molecular evolution is to calculate the ratio of synonymous (selectively unimportant) to nonsynonymous (selectively important) substitutions (Ina 1995; Yang and Nieson 2000). A high proportion of nonsynonymous substitutions indicate that a gene has been under positive selection (Messier & Stewart, 1997). This has been shown to apply to snake venom phosphorlipase molecules under selection pressure of resistant prey/predators (Nakashima *et al*. 1993). However most examples of positive selection inferred from nucleotide sequences involve host pathogen interaction as part of

evolutionary arms races between pathogens and their prey (Endo. *et al.* 1996). Bishop *et al.* (2000) report that dN in plant chitinase often exceeds dS rates under selection pressure for resistance to fungal pathogens.

Consequently, gene duplication and divergence play an important role in generating the functional diversification necessary for adaptation (Graur and Li, 2000). However, the contribution gene families make to ecological diversification and the nature of the evolutionary forces acting during this process remain poorly known. In part this is because genes directly involved in ecological attributes are hard to identify (Duda and Palumbi, 1999).

Fungal protease genes should provide a good model for studies of adaptive evolution of multigene families. The majority are secreted directly into the environment, and so can *prima facie* be defined as ecological traits, their production is repressed by readily utilized nutrients but induced by proteins and there are clear selectable differences in their substrate specificities (Murphy and Walton, 1996, Rawlings and Barrett, 1995, Siezen and Leunissen, 1997, St. Leger, Cooper and Charnley, 1987). Thus gene divergence and expression in genetically and ecologically distinct species can be more readily interpreted in terms of adaptation. For example, increases and decreases in protein family sizes in fungi could correlate with differences in function that are indicative of adaptations to environment and life strategies. Expressed sequence tag (EST) analyses showed that a broad host range strain of *M. anisopliae* sf. *anisopliae* (strain 2575) expressed 11 subtilisins during growth on insect cuticle, the largest number of subtilisins reported from any fungus. Polymerase chain reaction amplified 10 of their orthologs from a second

strain with multiple hosts (strain 820) and seven from the locust specialist *M. anisopliae sf. acridum* (strain 324). Analyses based on sequence similarities and exon–intron structure grouped *M. anisopliae* subtilisins into four clusters—a class I ("bacterial") subtilisin (Pr1C), and three clusters of proteinase K-like class II subtilisins: extracellular subfamily 1 (Pr1A, Pr1B, Pr1G, Pr1I and Pr1K), extracellular subfamily 2 (Pr1D, Pr1E, Pr1F and Pr1J) and an endocellular subtilisin (Pr1H). Phylogenetic analysis of homologous sequences from other genera revealed that this subdivision of proteinase K-like subtilisins into three subfamilies preceded speciation of major fungal lineages. However, diversification has continued during the evolution of *Metarhizium* subtilisins with evidence of gene duplication events after divergence of *M. anisopliae* sf. *anisopliae* and *M. anisopliae* sf. *acridum*. Comparing alignments and nonsynonymous/synonymous rates for Pr1 isoenzymes within a lineage and between lineages showed that while overall divergence of subtilisins followed neutral expectations, amino acids involved in catalysis were under strong selective constraint. This suggests that each Pr1 paralog contributes to the pathogens fitness. Furthermore, homology modeling predicted differences between the Pr1's in their secondary substrate specificities, adsorption properties to cuticle and alkaline stability, indicative of functional differences. While it is not clear which protease is responsible for any given activity, re scavenging for nutrients or pathogenicity related, strong selective constraint against mutations in active site regions of each enzyme confirms that they are not redundant and each must have a role (Bagga *et al*., 2004).

Chapter 2: A phylogenomic approach to reconstructing the diversification of serine proteases in fungi

Introduction

The serine endoproteases are divided into two superfamilies that independently evolved similar catalytic mechanisms. The trypsin superfamily includes the trypsins and chymotrypsins that are ubiquitous in animals. The subtilase superfamily is similarly ubiquitous in bacteria and fungi. The classification of Siezen & Leunissen (1997) delineates three families of subtilases in fungi, the proteinase K family (fungal or class II) named after an enzyme found in the ascomycete *Tritirachium album*, the subtilisin family (also called bacterial or class 1) which until recently were only thought to be found in bacteria, and kexins (preprotein convertases) that are also found in animals.

Fungi depend for their life activities on their ability to harvest nutrients from living or dead plant and animal material. The ecological diversification of fungi is therefore profoundly affected by the array of enzymes they secrete. Thus, fungi pathogenic to plants and animals show adaptation in the range of enzymes produced to the polymers present in the integuments of their particular hosts; the ability to colonize these taxa among different fungal species being linked to enzyme evolution (St. Leger *et al*., 1997). Among many saprophytes, subtilases are the principal broad-spectrum proteases (Gunkle & Gassen, 1989). Their ubiquity among fungi suggests that they are unlikely to be specifically developed to implement pathogenicity. In fact, during the evolution of many plant pathogens the functions of a broad-spectrum protease are performed by trypsins (Murphy & Walton, 1996; St.

Leger *et al*., 1997; Bidochka *et al*., 1999). Transcripts of two trypsin genes are also the most highly expressed by the insect pathogen *Metarhizium anisopliae* under some conditions (Freimoser *et al*., 2003a).

Most fungal genes encoding hydrolytic enzymes are members of gene families-genes of common origin that encode products of similar function (Walton, 1996). However, the contribution gene families make to ecological diversification and the nature of the evolutionary forces acting during this process remain poorly known. In part this is because genes directly involved in ecological attributes are hard to identify (Duda & Palumbi, 1999).

Fungal protease genes should provide a good model for studies of adaptive evolution of multigene families. The majority are secreted directly into the environment, and so can prima facie be defined as ecological traits, their production is repressed by readily utilized nutrients but induced by proteins and there are clear selectable differences in their substrate specificities (St. Leger *et al*., 1987; Rawlings & Barrett, 1995; Murphy & Walton, 1996; Siezen & Leunissen, 1997). Thus gene divergence and expression in genetically and ecologically distinct species can be more readily interpreted in terms of adaptation. For example, increases and decreases in protein family sizes in fungi could correlate with differences in function that are indicative of adaptations to environment and life strategies. *Metarhizium anisopliae* produces at least 11 subtilisins; to date the largest number reported from any fungus (Bagga *et al*., 2004) but only two trypsins (Freimoser *et al*., 2003a). Although it is not clear which specific protease is responsible for any given activity, either scavenging for nutrients or pathogenicity related, strong

selective constraint against mutations in active site regions of each enzyme confirms that they are not redundant and each must have a role (Bagga *et al*., 2004).

Sequences of subtilisin and trypsin homologues continue to pour into sequence databases, including from genome projects. At the time this work was submitted there were nine virtually completed sequenced fungal genomes representing six ascomycete and three basidiomycete lineages [Goffeau *et al*., 1996; Wood *et al*., 2002; Galagan *et al*., 2003, Genome sequencing projects for *Aspergillus nidulans, Coprinus cinereus, Cryptococcus neoformans, Fusarium graminearum, Magnaporthe grisea, Neurospora crassa* (version 3) and *Ustilago maydis*, Center for Genome Research (http://www.broad.mit.edu)]. These fungi have very different virulence and habitat that provide a formidable new resource for determining the identity, origin and evolution of traits that contribute significantly to fitness.

In this study I used genomic sequences and expressed sequence tag (EST)/polymerase chain reaction (PCR) data from a project designed to identify the full spectrum of proteases in *M. anisopliae* (Freimoser *et al*., 2003a; Bagga *et al*., 2004), to identify likely evolutionary events such as patterns of gene duplications and gene loss in the history of the subtilase and trypsin superfamilies in fungi. These events would likely reflect specialization and adaptation. The phylogenomics approach, combining phylogenetic reconstruction and analysis of complete genomic sequences (Eisen & Fraser, 2003), allows estimates of the age of duplication events, which in turn can greatly aid in functional studies (i.e. recent duplications suggest the expansion of an activity in a species, old duplications likely reflect divergent

functions) (Heidelberg *et al*., 2000; Jordan *et al*., 2001). In addition, in order to reconstruct trypsin evolution, 35 representative fungi were use to detect the presence or absence of trypsin genes. This is because available sequence data derives from too few fungal species to conduct a comparative analysis on the evolution of gene diversity. The origin of the trypsin superfamily in prokaryotes or eukaryotes is controversial (Rawlings & Barrett, 1995), as until this study trypsin homologues had only been found in actinomycete bacteria, five pathogenic ascomycetes, including *M. anisopliae*, and animals.

This study demonstrated that a substantial diversification of subtilase-type proteases occurred early in ascomycete history (with subsequent loss in saprophytic lineages). However, the pathogens retained and occasionally expanded different gene families. Thus, *M. grisea* has 15 subtilisins and six proteinase K subtilases, whereas *M. anisopliae* and *Fusarium graminearium* each possess 11 proteinase K subtilases but three or fewer subtilisins. Trypsin genes are lacking in most saprophytes, but are present in a basidiomycete insect symbiont (*Septobasidium canescens*), most zygomycetes and many ascomycete plant and insect pathogens. The patchy distribution of trypsins suggests that their phylogenetic breadth will have been much wider in early fungi than currently.

Materials and methods

Organisms

Allomyces macrogynus (ATCC 38327), *Mucor mucedo* (ATCC 38694), *S. canescens* (ATCC 20021), *Sporobolomyces roseus* (ATCC 24257), *C. cinereus* (ATCC 20120) and *Leptosphaeria taiwanensis* (ATCC 38203) were obtained from

the American type Culture Collection. *Conidiobolus coronatus* (ARSEF512), *C. lamprauges* (ARSEF), *C. obscurus* (ARSEF133), *Neozygites parvispora* (ARSEF320), *Zoophthora radicans* (ARSEF1341), *Basidiobolus ranarum* (ARSEF264), *Paecilomyces fumosoroseus* (ARSEF5540), *Hirsutella thompsonii* (ARSEF194) and *M. anisopliae* sf. *anisopliae* (ARSEF2575) were obtained from the US Department of Agriculture Entomopathogenic Fungus Collection in Ithaca, NY *Rhizopus stolonifer* (ER-15-6223), *Penicillium chrysogenum* (=*Penicillium notatum*) (ER-15-6157), *Sphaerostilbella lutea* (GJS 82-274), *Hypomyces chrysospermus* (GJS 97-173), *Hypocrea gelatinosa* (CBS 887-72), *Diaporthe arctii* (AR2831) and *Bionectria ochroleuca* (GJS90-167) were obtained from Dr Mary Rossman at the US Department of Agriculture Fungus Collection in Beltsville, MD, USA. Cultures were maintained on Sabouraud dextrose agar (DIFCO, Franklin Lakes, NJ, USA). *Agaricus bisporus* mushrooms were obtained from a local market.

Trypsin sequences from fungi

Mushrooms (*A. bisporus*) were used directly as a source of genomic DNA. Otherwise, genomic DNA from 2-day-old Sabouraud dextrose broth cultures of each fungus was isolated as described (Screen & St. Leger, 2000). Degenerate primers based on an alignment of all available fungal trypsins were used as templates for PCR. They were forward GTISTIACYGCNGSYCAYTG and reverse AKIGGRCCICCRDWRTCDCC (where I = Inosine; S = C,G; Y = C,T; N = A,G,C,T; K = G,T; R = A,G; D = A,G,T; W = A,T). Standard cycling conditions following optimization were 94 °C for 4 min, one cycle; 94 °C for 1 min, 4548 °C for 1 min, 72 °C for 1 min, 32 cycles; 72 °C for 7 min, one cycle. PCR reactions

were analyzed on gels, the bands of interest excised, purified with a Qiagen gel extraction kit (Qiagens Science, Germantown, MD, USA), and the PCR products sequenced (see Fig. 4 for accession numbers). To confirm the absence of trypsins, we probed Southern blots with a mixture of trypsin fragments from *M. anisopliae, C. coronatus* and *S. canescens*. Blots were performed at moderate stringency using standard procedures (St. Leger *et al*., 1992).

Phylogenetic analysis

The sequences of previously characterized trypsins and subtilases were downloaded from the National Center for Biotechnology Information (NCBI) databases using *M. anisopliae* sequences as query sequences in blast, blast2 and PSI-blast searches (accession numbers or genomic locus ID numbers are given in Figs 1-4). Databases searched included genome sequences of three Basidiomycetes, *C. neoformans* (an opportunistic pathogen of immunocompromised individuals) C. cinereus (saprophyte), *U. maydis* (plant pathogen), two saprophytic ascomycetes (*A. nidulans, N. crassa*) and two plant pathogenic ascomycetes (*M. grisea, F. graminearum*) at the Whitehead Institute. Databases for *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* were accessed at Stanford University and the Sanger Institute, respectively. DNA and protein sequence alignments were generated with Clustal W (Thompson *et al*., 1994). Signal peptides were predicted using SignalP V2.0.b2 (Brunak & Von Heijne, 1999) at http://www.cbs.dtu.dk/services/SignalP-2.0. Exons were predicted by alignment with known subtilsins and using the GT/AG rule. Phylogenetic analyses on the aligned amino acid sequences were performed using the neighbour-joining method

(Poisson correction) in the MEGA 2.1 (Kumar *et al*., 2001) program. Confidence in the topology of phylogenetic trees was evaluated by performing 1000 bootstrap replicates in the program. Comparisons of sequence pairs in the alignments for each gene group were made by calculating synonymous (d S) and nonsynonymous (d N) nucleotide substitutions with the Yang *et al*. (2000) model for codon change in the PAML program (Department of Biology, UCL, London, UK).

Results

Phylogenetic analysis of the subtilase superfamily ubsection

Three families of fungal subtilases (Siezen & Leunissen (1997), the proteinase K family (fungal or class II) (Fig. 1), the subtilisin family (bacterial or class 1) (Fig. 2), and kexins (preprotein convertases) (Fig. 3) were included in this study. We placed these families in separate phylogenies, as they are very divergent from each other. Throughout the paper the ascomycete subtilisin subfamilies are abbreviated as SF1 and SF2 to delineate them from the proteinase K subfamilies (sf1 and sf2).

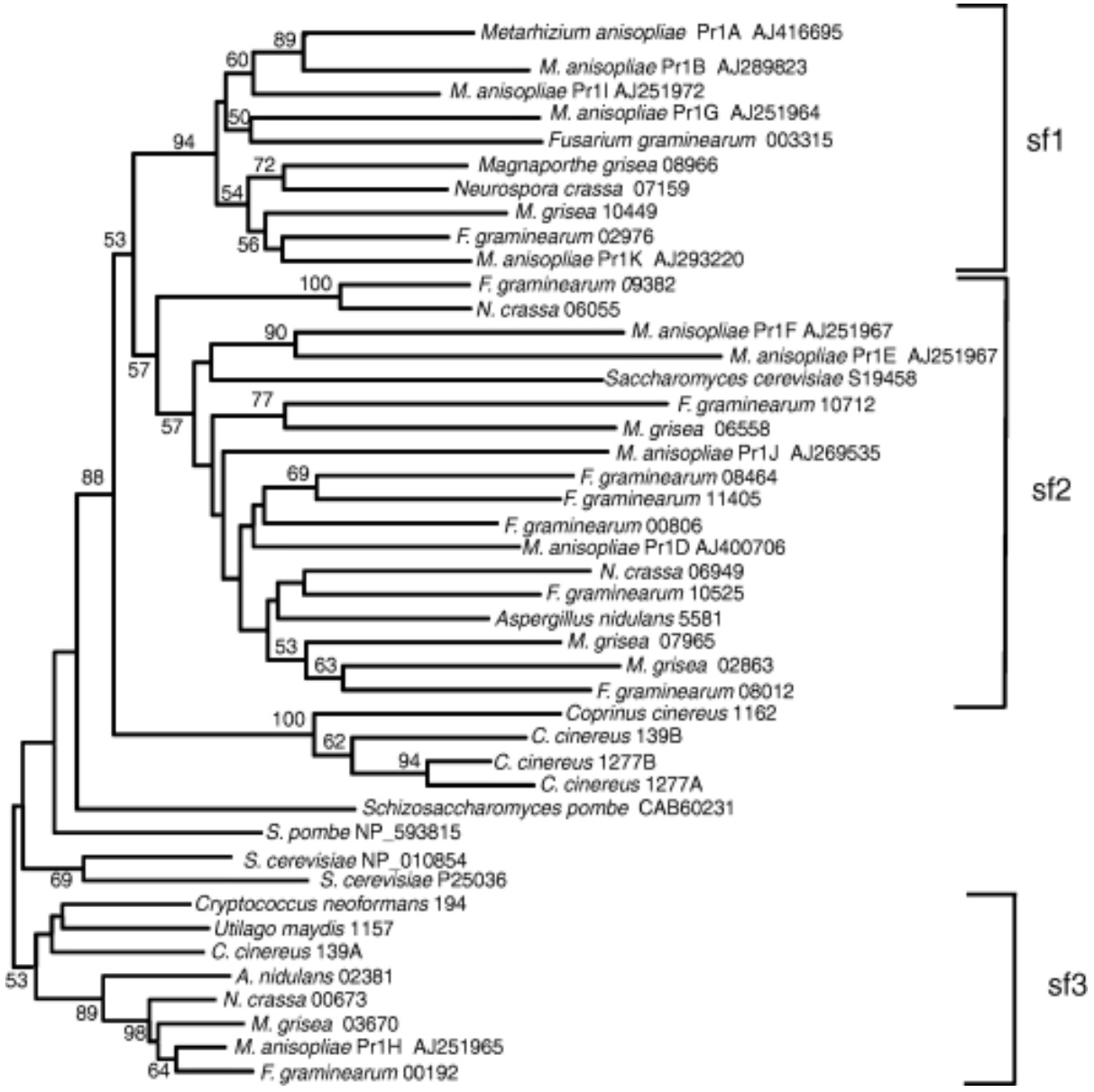
Proteinase K family

A previous survey of published subtilases identified three distinct clusters within the proteinase K family in fungi, two subfamilies containing only secreted proteases (sfs1 and 2) and a subfamily that includes endocellular subtilisins (sf3) (Bagga *et al.*, 2004). A similar topology was found in the gene sequences in genomic databases, albeit sf2 receiving weaker bootstrap support (60%) than sf1 (93%). The sf3 group was without bootstrap support when *S. pombe* sequences were

included, but in their absence the *S. cerevisiae* sequences formed a cluster with the sf3 sequences that received 96% support (data not shown). The presence of at least one sf3 proteinase K like enzyme (Fig. 1) and a bacterial class 1 subtilisin (Fig. 2) in each of the basidiomycetes demonstrates that the last common ancestor before the basidiomycete/ascomycete split [*ca*. 500 Ma (Berbee & Taylor, 2001)] possessed proteinase K and subtilisin proteases. However, the four proteinase-K subtilases of the basidiomycete *C. cinereus* are basal to the sf1/sf2 split and cluster together (100% bootstrap support) suggesting that the sf1-sf2 duplication occurred in the Ascomycota. This presumably preceded the filamentous ascomycete/yeast split approximately 400 Ma (Berbee & Taylor, 2001) as *S. cerevisiae* has a single sf2 enzyme.

At first sight, by the measure of total number of sequences, gene duplication would seem to have occurred considerably more often in pathogenic ascomycetes than in the other groups. However, the majority of sf1 and sf2 subtilases from *M. anisopliae*, *F. graminearium*, *M. grisea* are dispersed throughout the assemblage of sequences indicating that they originated from a common ancestor and that many duplication events preceded speciation of ascomycete lineages. Divergence of these lineages may have been taking place 240 Ma (Berbee & Taylor, 2001). Other sequences clustered with sequences from the same species, suggesting more recent duplications. Consistent with these being recent duplications the lineage specific clusters consist of just two or three genes e.g. *F. graminearium* sequences 1467 and 1340, and *M. anisopliae Pr1A*, *Pr1B* and *Pr1I*.

Fig. 1 Neighbour-joining tree of proteinase K gene family protein sequences. Accession numbers for S. cerevisiae, and S. pombe, locus ID (*A. nidulans , F. graminearium, M. grisea and N. crassa*) or contig numbers (*C. neoformans, C. cinereus and U. maydis*) as available from the Whitehead Institute web site (http://www-genome.wi.mit.edu/) are provided in the figure. Accession numbers for *M. anisopliae* Pr1A to Pr1K are from Bagga *et al*. (2004). Genes classifiable into subfamilies 1 (sf1), 2 (sf2) and 3 (sf3) and bootstrap support (>50%) are also shown. Full names of each species are given in the first instance in the figure.



In a minority of cases it was possible to assign orthologues. This was most straightforward with the three genes from *N. crassa*. Thus *N. crassa* sequence 07159 with *M. grisea* sequence 08966; *N. crassa* sequence 06066 with *F. graminearium* sequence 1383 and *N. crassa* sequence 06949 with *F. graminearium* sequence 1441 (albeit the last one with low bootstrap support).

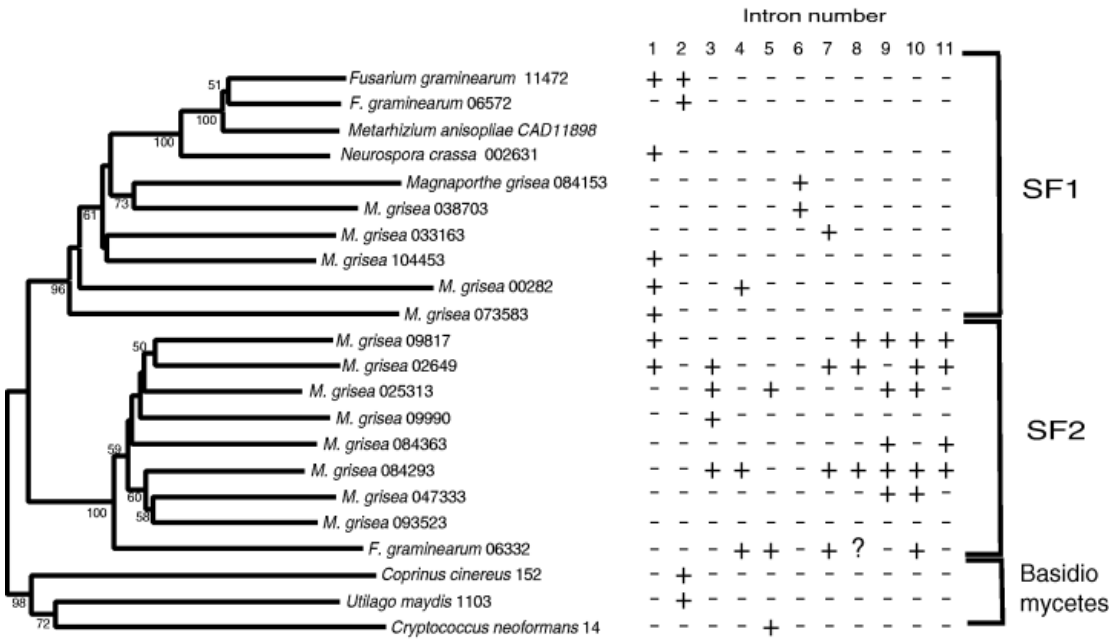
The class I subtilisin family

Aspergillus nidulans, *S. pombe* and *S. cerevisiae* have lost subtilisins. Otherwise the genomes contain a single subtilisin gene except for *F. graminearium* and the rice blast fungus *M. grisea* that contain three and 15, respectively. All the sequences predict a signal peptide suggesting they are secreted. One of the *M. grisea* genes is truncated and therefore not considered in the phylogeny (Fig. 2).

The tree contains three clusters; a basidiomycete grouping and two subfamilies of ascomycete sequences (>96% bootstrap support), each containing some *M. grisea* genes (ascomycete subtilisin subfamilies are abbreviated as SF1 and SF2 to delineate them from the proteinase K subfamilies, sf1 and sf2). Subfamily 1 was less tightly clustered than SF2 and includes a subcluster of sequences from *M. anisopliae*, *N. crassa* and *F. graminearium*. The tree provides evidence for an independent duplication in *Fusarium*, as the two SF1 *F. graminearium* sequences, Fg11472 and Fg506572, cluster together with relatively high pairwise similarity (59%). Because SF2 also contains a *F. graminearium* sequence the tree indicates that the SF1-SF2 duplication preceded radiation of the ascomycete orders, Sordariales (*M. grisea* and *N. crassa*) and Hypocreales (*F. graminearium* and *M.*

anisopliae). We can also infer from their absence the differential loss of SF2 genes in *N. crassa* and possibly *M. anisopliae* (inferred from failure to find an EST).

Fig. 2 Neighbour-joining tree of subtilisin class I gene family protein sequences. Bootstrap support (>50%), locus ID (*F. graminearium, M. grisea, N. crassa*) or contig numbers (*C. cinereus, C. neoformans and U. maydis*) (from the Whitehead Institute web site) and accession numbers (for *M. anisopliae*) are provided in the figure. The 11 intron positions shared by two or more genes are numbered consecutively from 5' to 3'. Genes classifiable into subfamilies 1 (SF1) and 2 (SF2), and the subfamily of basidiomycete sequences are also shown.



Nevertheless, as most basidiomycetes and ascomycetes possess just one subtilisin, the parsimonious hypothesis based on the data would be that the multigene family in *M. grisea* originated primarily through duplication at the origin of the genera. Consistent with this, most of the *M. grisea* sequences clustered with

other *M. grisea* sequences. However, the average percent sequence identity (range) between *M. grisea* subtilisins in SF1 and SF2 are only 23% (1833) and 36% (3045), respectively, which suggests divergence early in ascomycete history. This conclusion is based on the lower level of sequence divergence between Fg11472 and Fg506572 in *F. graminearium* and the similarity of these sequences to their homologues in *M. anisopliae* (55%) and *N. crassa* (48%). In addition, although the six proteinase K *M. grisea* subtilases are dispersed among sf1, sf2 and sf3 sequences from other genera (Fig. 1) they show similar sequence divergence [36 (2558)%] to the subtilisins suggesting they duplicated over a similar time scale. In contrast, the four proteinase K genes in C. cinereus (Fig. 1) appear to have diverged comparatively recently, and perhaps within the Coprinus genus itself, judging by the comparatively high sequence similarity between them [54 (4861)%] and the presence of the two most similar sequences (Cc1277A and Cc1277B) 6000 base pairs apart on the same scaffold, consistent with tandem duplication. In contrast, the two closest *M. grisea* sequences were >20 000 base pairs apart. Likewise the three *N. crassa* proteinase K sequences have average identities of 46% (3767) with their putative orthologues. Consistent with both fungi belonging to the same order, the *N. crassa* 07159/*M. grisea* 08966 sequences are 67% similar. Thus, sequences with <46% identity may have diverged before the origin of major ascomycete lineages. It is also likely that purifying selection for functionally important domains will have constrained divergence much below the 35% sequence similarity found in the subtilisins.

Recently duplicated genes could show accelerated evolution by being free to accumulate amino-acid changes toward functional divergence (Graur & Li, 2000). We calculated d N/d S ratios (the ratio of amino acid altering substitutions to silent substitutions), to determine, first, whether d N/d S varies between the subtilisin and proteinase K multigene families in *M. grisea* and, secondly, whether positive selection governs gene family diversification. The *M. grisea* subtilisin gene grouping shows an average pairwise d N/d S of 0.742/0.976 = 0.76 ± 0.07. This is similar to the diversification in the proteinase K genes (0.717/0.825 = 0.903 ± 0.08), and provides no evidence for increased selection in the subtilisins.

A further indication that the duplications are not recent is that the d S values for subtilisin genes are close to 1, i.e. divergence has been sufficient to allow saturation of nucleotide substitutions. At this level of divergence it may be difficult to detect positive selection among duplicated genes, simply because positive selection may only occur in a short evolutionary time after gene duplication during the functional shift of the protein, and its effect can be obscured by later substitutions (Zhang *et al*., 1998).

Structure of *M. grisea* subtilisin genes

Detection of intron losses is a useful tool for interpreting the origin, development, and divergence of multigene families, and for reconstruction of gene evolution (Frugoli *et al.*, 1998). Consistent with *M. grisea* subtilisins not being recent duplications, their genetic structure was very different as shown by mapping intron insertion sites on the multiple protein sequence alignments. The identities of intron locations were considered stringently. Thus, Fig. 2 shows introns located at

the same position in the alignment and with the same phase in the codon. We found only one pair of closely separated (four codons apart) intron sites among genes (indicated by a question mark in Fig. 2). Thus, if intron sliding occurs at all, it must be much rarer than loss or gain of introns.

Two introns in basidiomycete sequences had counterparts in some ascomycete genes. Intron 2 is common to *F. graminearium* sequences 11472 and 06572 (SF1), and to homologues in the basidiomycetes *C. cinereus* (*Cc152*) and *U. maydis* (*Um1103*). Intron 5 is common to SF2 genes *Mg025313* and *Fg06332*, and the basidiomycete *C. neoformans* gene, *Cn14*. Given that the shared basidiomycete/ascomycete introns are in either SF1 and SF2, and the basidiomycete sequences cluster together (98% bootstrap support) rather than with SF1 or SF2, the simplest explanation for the SF1/SF2 split is that the duplication that gave rise to the progenitors of SF1 and SF2 predated the divergence of basidiomycetes and ascomycetes from a common ancestor. In which case, the ancestral ascomycete inherited two subtilisins that possessed introns 2 and 5, respectively. However, we can also infer from the discontinuity between basidiomycete and ascomycete sequences that most duplication events within SF1 and SF2 occurred after the basidiomycete/ascomycete split. Introns 1, 4 and 7 have counterparts in some sequences from both SF1 and SF2. As they are present in *F. graminearium* as well as *M. grisea*, they predate speciation within ascomycetes, confirming some duplication events within both SF1 and SF2 preceding radiation of ascomycete orders with subsequent differential loss of genes in different lineages.

There is extensive evidence for multiple intron loss in the phylogeny. Intron 2, being present only in *F. graminearium* and basidiomycete genes was presumably lost early in the lineage leading to *M. grisea*, before the divergence of most sequences, whereas the patchy distribution of introns 1, 4, 5 and 7 provide examples of multiple independent instances of intron loss. Inspite of losses of single introns such as intron 2 in *Fg*06572 and intron 10 in *Mg*084363, most genes with few shared introns (e.g. *Fg*11472, *Fg*06572, *Nc*002631, *Mg*104453, *Mg*00282, *Mg*073583, *Cc*152, *Um*1103) tend to have those introns at the 5' termini. Also, in each case, the intron losses were exact without changes in the surrounding codon sequence. These circumstances favour a model for concerted loss of introns through gene conversion with intronless (cDNA) copies of the gene (Hartung *et al.*, 2002). An exception, *Mg047333*, only retains introns 10 and 11 and so the recombination break point may have lain upstream of intron 10. This mechanism for concerted intron loss may have been more active in some clusters than other, as sf1 paralogues have fewer introns than sequences in sf2.

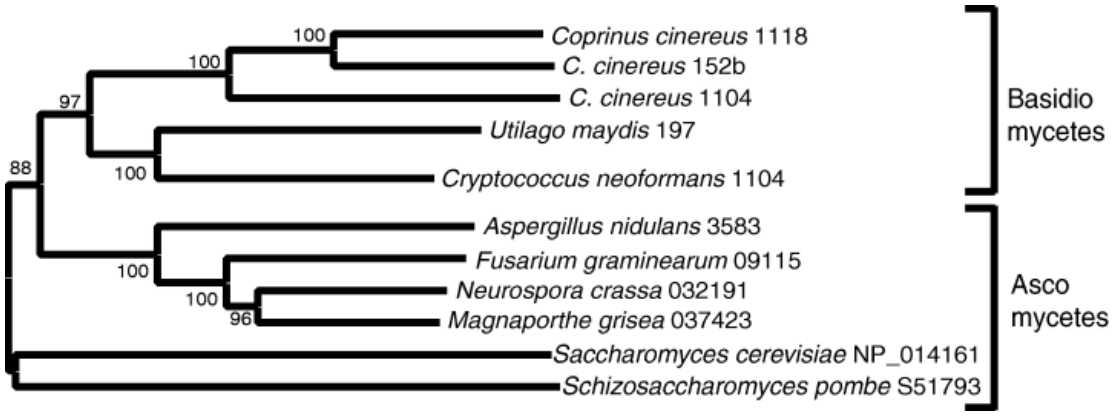
Following events of intron loss, the appearance of a new intron in a novel position is consistent with a mechanism of intron gain because one gain event is more likely than many independent intron loss events (Dibb & Newman, 1989; Hartung *et al.*, 2002). Most paralogues had one or more unique introns that cannot be aligned clearly to any other gene. This is consistent with ancient gene duplication events, as sufficient time must have elapsed since duplication to allow for intron additions. The exclusive presence of intron 6 in the sub-cluster *Mg084153* and *Mg038703* also suggests recent gain. However, intron loss seems common and

widespread in the subtilisins, which emphasizes the difficulty of reconstructing the evolution of gene structure because of a labile nature of intron presence or absence (Krzywinski & Besansky, 2002).

Kexins

The tree topology recovered from the subtilisin data likely reflects a combined effect of phylogenetic relationships, and uneven patterns of gene duplication and loss. A much simpler case is presented by the fungal kexins, probably because they appear to have diversified principally in animals. We searched the human genome for kexin-like sequences and identified 13 distinct kexin genes. In contrast, individual fungi contain three at the most (Fig. 3). All the sequences predicted an endocellular function consistent with their role as preprotein convertases (Siezen & Leunissen, 1997). There is evidence from degree of divergence for relatively recent duplication of the kexin gene in *C. neoformans* and more ancient duplication events in the lineage leading to *C. cinerea*. However, aside the anomalous basal position of yeast sequences the observed phylogenetic distribution of the kexin orthologues is in agreement with taxonomically accepted relationships with excellent statistical support at most nodes (Fig. 3).

Fig. 3 Neighbour-joining tree of kexin-like subtilisin protein sequences. Bootstrap support (>50%), locus ID (*A. nidulans, F. graminearium, M. grisea* and *N. crassa*) or contig numbers (*C. cinereus, C. neoformans* and *U. maydis*) (from the Whitehead Institute web site), and accession numbers (*S. cerevisiae*, and *S. pombe*) are provided in the figure.



Phylogenic analysis of the trypsin superfamily

In order to reconstruct trypsin evolution in fungi and investigate the possibility that gene retention may represent niche-specific traits (traits shared by organisms that occupy the same niche irrespective of their phylogenetic position), we screened for the presence or absence of trypsins in 35 representative fungi (Fig. 4). To confirm the absence of trypsins, we probed Southern blots with trypsin genes (Fig. 5). In 33 cases PCR and Southern analyses were consistent. However, bands were detected in DNA blots of the saprophytic zygomycete *M. mucedo*, the plant pathogenic ascomycete *B. ochroleuca* and the basidiomycete phylloplane/opportunistic pathogen *S. roseus*, although PCR reactions using their

DNA as template did not produce trypsin fragments. One possibility is that these represent degenerating sequences not recognized by primers as we previously failed to detect trypsin activities in *Mucor* spp. (Freimoser *et al.*, 2003b). Although failure to detect trypsins using PCR technology/Southern analysis does not prove loss, the observed cases are totally consistent with available genomic data. Thus, PCR successfully amplified the trypsin sequences from *A. nidulans* and *F. graminearum* although not detecting trypsins in the other sequenced fungi, and this interpretation was supported by Southern analysis.

Trypsin genes were not detected in five of the seven basidiomycetes screened. The exceptions, clustering together in the phylogenetic tree (Fig. 4) were *S. roseus* (Southern blot data only) and *S. canescens*, an obligate scale insect symbiont that penetrates (using proteases?) its hosts cuticle. Trypsins were also present in six of 11 ascomycete pathogens of plants and insects, and the mycopathogen *T. harzianum*. *A. nidulans* was the only one of eight saprophytic ascomycetes with a trypsin. Trypsin genes were detected in six of eight zygomycetes, all but one a pathogen, showing that their last common ancestor with the basidiomycete/ascomycete clade had this enzyme [about 850 Ma (Berbee & Taylor, 2001)]. We were also able to find a trypsin in *B. ranarum*, an increasingly important pathogen of animals and humans that may be a chytrid (Bruun *et al.*, 1998). However, no trypsin was detected in the basal chytrid *A. macrogynus*.

Fig. 4 (a) A consensus tree showing the relationships among 35 fungi as inferred from 18S rRNA gene sequence data. Bootstrap support (>50%) is provided in the figure. Taxa name in bold means that we PCR amplified at least one trypsin. Asterisks denote validation of the presence of trypsins by genome sequencing. Taxa name underlined means that the genome of the fungus has been sequenced but contains no trypsin genes. Taxa name in uppercase means one or more bands cross-hybridizing with trypsin genes in Southern blot analysis although PCR reactions did not produce trypsin fragments (Fig. 5). (b) Neighbour-joining tree of trypsin proteins showing congruency with the 18s RNA sequences. Bootstrap support (>50%) and accession numbers are provided in the figure. Sequences from *Fusarium oxysporum*, *Cochliobolus carbonum, Trichoderma harzianum* and *Phaeospaeria nodorum* were obtained from public data bases. The accession numbers for the 18s RNA sequences used in (a) are: A. bisporus, AJ244527; *A. macrogynus*, U23936; *A. nidulans*, X78539; *B. ranarum*, AF113414; *B. ochroleuca*, AH007787; *C. carbonum*, U42479; C. coronatus, AF113417; C. lamprauges, AF296754; C. obscurus, AF368508; *C. coprinus*, M92991; *C. neoformans,* M55625; *D. arctii*, L36985; *F. oxysporum*, Z94126; *H. thompsonii*, U32406; *H. gelatinosa*, U32407; *H. chrysospermus*, M89993; *L. taiwanesis*, U43447; *M. grisea*, AF277124; *M. anisopliae*, AF280631; *M. mucedo*, X89434; *N. parvispora*, AF296760; *N. crassa*, AY046271; *P. fumoso-roseus*, AB032475; *P. chrysogenum* (=*P. notatum*), L76153; *P. nodorum*, U04236; *R. stolonifer*, AF113441; *S. cerevisiae*, AF331938; *S. pombe*, AY046272; *S. canescens*, AY123320; *S. lutea*, U32415; *S. roseus*, X60181; *T. harzianum*, AF548100; *U. maydis*, X62396; *Z. radicans*, D61381. The 18s RNA sequence for *F. graminearum* was obtained from the *F. graminearum* excluded reads database at Whitehead Institute web site (http://www-genome.wi.mit.edu/), sequence G578P61595RER (1051114).

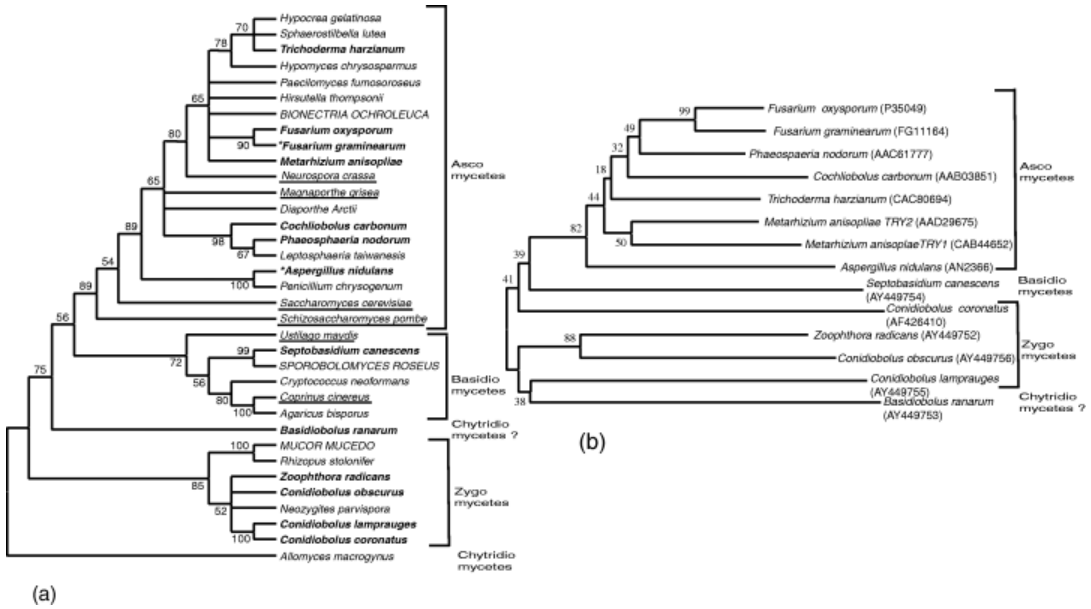
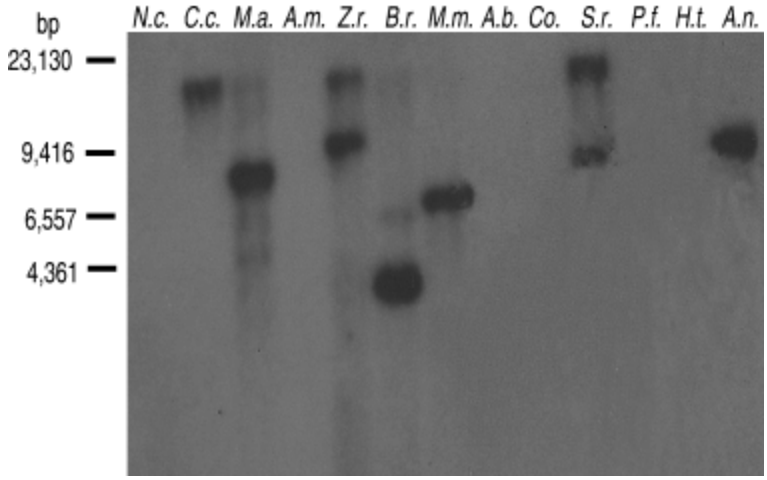


Fig. 5 Southern blot analysis of representative fungi probed with PCR amplified trypsin fragments from *M. anisopliae, C. coronatus and S. canescens. Abbreviations are: N.c., N. crassa; C.c., C. coronatus; M.a., M. anisopliae; A.m., A. macrogynus; Z.r., Z. radicans; B.r., B. ranarum; M.m., M. mucedo; A.b., A. bisporus; Co., C. cinereus; S.r., S. roseus; P.f., P. fumosoroseus; H.t., H. thompsonii; A.n., A. nidulans.*



The estimate of phylogeny obtained using trypsin sequences (Fig. 4b) is largely congruent with the unconstrained rDNA tree (Fig. 4a), which is consistent with current estimates of organismal phylogeny for the fungi (Berbee & Taylor, 2001). This suggests a pathway of evolution in which the divergence of trypsins largely reflects the speciation of fungal lineages and the absence of trypsins in specific fungi reflects gene loss. Although the bootstrap support for many groups in the trypsin tree is low, these low values simply emphasize the absence of conflict with the fungal species tree.

Discussion

The present study has revealed evolutionary selection of one gene family of proteases over another in species of ascomycete fungi. This suggests that differences in the properties of the enzymes provide selective advantages in different habitats. Although we presently cannot link these observations to any organisms' lifestyle in specific terms, the presence of large clusters of genes is likely to reflect selective pressure for their increased or varied coding capacity (Jordan *et al.*, 2001). The multiple subtilases in pathogenic ascomycetes could play different roles in pathogenesis, increase adaptability and host range, or have different functions in survival in various ecological habitats outside the host. There are clear catalytic differences between subtilases and trypsins, and within the proteinase K family of subtilases in their substrate specificities (St. Leger *et al.*, 1994, 1996) that could allow their selection. In the case of *M. anisopliae* proteinase K subtilases (Pr1 enzymes) this extends to probable differences in their stability, absorption properties to insoluble substrates and interactions with protease inhibitors (Bagga *et al.*, 2004). The selectable properties of class I subtilisins as compared to proteinase K subtilases is unknown. It remains to be determined if diversity in function parallels the subtilisin gene family diversity in *M. grisea* that exceeds that found in other fungi. The unusual number of subtilisins implies some particular adaptive significance for this lineage. Possibly the retention by *M. grisea* of so many subtilisin paralogues simply results from selection to achieve high expression by gene dosage; an example of gene copy number selection (Gruhl *et al.*, 1997). This

requires experimental validation to determine if the isoenzymes encoded by proteinase K genes change more in response to environments than those encoded by subtilisin genes, suggesting their stronger regulation (expression).

The divergence of subtilisins seen in *M. grisea* may be sufficiently ancient as to represent the ancestral condition in early ascomycetes. The results, both from protein sequence data and intron profiling, are consistent with most duplication events occurring after the ascomycete/basidiomycete split but before radiation of major fungal lineages such as the Sordariales and Hypocreales. If this is the case, then the relative paucity of these genes in most lineages reflects gene loss as compared with *M. grisea*. Uncertainty surrounding the timing of duplication and loss events will remain problematic, pending improved resolution of basal lineages of ascomycetes and extended taxon sampling. Clearly however, the absence of subtilisins in three distantly related fungi (*A. nidulans, S. pombe* and *S. cerevisiae*) is consistent with a trend towards loss of this group in diverse lineages.

Every fungus in this study possessed proteinase K subtilase genes. That includes the protease-impoverished yeast species that lack subtilisins but contain at least one proteinase K gene. Comparing the yeast genomes to those of other fungi suggests gene loss has taken place at many different points during the evolution of the *S. cerevisiae* and *S. pombe* lineages (Braun *et al.*, 2000; Braun, 2003; Krylov *et al.*, 2003). However, one can imagine selection acting particularly strongly against the deletion of the last gene in a gene family. This suggests that the proteinase K subtilases are less expendable than subtilisins even in the yeasts carbohydrate rich, protein poor habitat. The high level of sequence polymorphism shown inside each

of the subtilase sub-families suggest that like subtilisins these genes also diversified in ascomycetes after the ascomycete/basidiomycete split but before radiation of several major ascomycete lineages. Aside the four proteinase K genes in *C. cinerea*, which appear to have diverged relatively recently, there is no evidence for a parallel divergence of genes encoding secreted proteases in basidiomycetes. The earliest basidiomycetes seem to have been wood inhabiting (Lewis, 1987), and in sharp contrast to the ascomycetes extant basidiomycetes include very few pathogens of animals, including insects (Carlile *et al.*, 2001). It is intriguing to speculate that extra proteolytic competence may have allowed the early ascomycetes to grow on a greater variety of living and nonliving proteinaceous substrates. This may have been a component allowing niche differentiation between the ascomycetes and the basidiomycetes that will have adapted the former to pathogenicity to animals or may itself have derived from adaptation to pathogenicity. In any event, the fact that two families of subtilases radiated in the early ascomycetes suggest that these fungi had a lifestyle that selected for multiple protease activities. That the majority of fungi have a single sf3 activity suggests that this selective pressure did not extend to endocellular housekeeping enzymes. *Aspergillus nidulans* as a Eurotiomycete (=plectomycetes) provides an interesting outgroup as these diverged about 400 Ma (Kasuga *et al.*, 2002) from the Sordariomycete (=pyrenomycetes) Ascomycota lineage that includes the Sordariales and Hypocreales. Aside the sf3 gene (*An02381*), *A. nidulans* has a single proteinase K subtilase (*An5581*) in sf2. Likewise, all genes cloned so far from *A. fumigatus* and *A. niger* fall within sf2 (Bagga *et al.*, 2004). However, related *Penicillium* spp (Eurotiomycetes) have sf1

genes (Bagga *et al.*, 2004) demonstrating that their last common ancestor with *Aspergillus* spp had sf1 genes too. Thus, as with subtilisins there has been a general trend for the proteinase K gene family to decrease in size on average for many lineages following the early diversification. The most obvious explanation is that as the requirements for proteases varied over time in different lineages so did the gene family. Selection for some lifestyles, particularly pathogenicity, may favour retention of large gene families of proteases whereas saprophytes such as *N. crassa* adapted to live principally on readily utilized carbohydrates have lost genes. In fact, for the pathogens there is evidence from small lineage specific clusters for more recent expansions in subsets of subtilases indicating that the genes in a family do not necessarily behave alike. Supporting this, evidence has been presented for recent tandem duplication of subtilase Pr1F as *M. anisopliae* sf *anisopliae* and *M. anisopliae* sf. *acridum* diverged (Bagga *et al.*, 2004). Clearly for the strongest and most general inferences, future comparisons should consist of matched pairs of closely related pathogens and nonpathogens replicated throughout the ascomycete tree.

As extended sampling, with more taxa representing basal fungal lineages is necessary to distinguish the contributions of gene loss and gain, we used this approach to reconstruct the evolution of trypsins. Overall, Fig. 4 suggests that independent loss of ancestral trypsins has occurred in many fungal lineages. Several species that have no trypsins are nested among species that do, indicating that the trypsin has been lost comparatively recently in the lineage leading to that species. For example, the trypsin in *A. nidulans* is without a counterpart in *P. notatum*. The

presence of a trypsin in *A. nidulans*, implies that it is not redundant in this species which may be linked with it lacking a subtilisin and having just two proteinase K activities. The convergence of the trypsin gene tree with the rDNA tree indicated that trypsins, where present, have diverged in parallel with the organisms in which they are expressed and were not obtained via horizontal gene transfer. If horizontal gene transfer is negligible then the evolutionary significance of lineage specific gene loss will be greater in fungi than in prokaryotes (Krylov *et al.*, 2003). The ecological implication of this is that enzymes lost by fungi are not regained making gene loss a one-way street that will reduce future adaptive options compared to genetically more diverse ancestors.

Like subtilases, the trypsins are inducible by environmental cues and secreted (St. Leger *et al.*, 1996). Their interactions with the environment presumably confer considerable selective functions in those fungi that express them at high levels and possibly pre-adapt them to certain lifestyles such as pathogenesis. The fact that both animal and plant host-associated genera harbour trypsins suggest that in some manner plants and animals offer similar physiological selective pressures. Conversely, in fungi that lost them, trypsins may have been rendered nonessential by functionally analogous subtilase genes, they may have otherwise ceased to provide any benefit or they may have actually interfered with ecological functions. The species tree (Fig. 4) has many exceptions to the correlation of trypsins and pathogenic lifestyle. Thus, *M. grisea* and the insect pathogen *H. thompsonii* lack trypsins, whereas *A. nidulans* possesses one. *Metarhizium anisopliae* may be a transitional species in this regards as some strains express high

levels of a broad spectrum trypsin whereas other strains produce lower levels of an enzyme with a narrow specificity for Phe-Val-Arg (St. Leger *et al.*, 1987; Bidochka *et al.*, 1999). In addition the trypsin gene appears to be silent in *M. anisopliae* sf. *acridum* strain 324 (Freimoser *et al.*, 2003a). Thus, there are several mechanisms available for different strains to adapt enzyme activities to their specific needs on their particular hosts. A plausible explanation for *M. anisopliae* sf. *acridum* strain 324 not expressing trypsin is that it has a very narrow host range and does not require a gene that facilitates opportunistic host jumping. Likewise, the obligate mite pathogen *N. parvispora* also lacks a trypsin although it clusters among other less fastidious zygomycetes that do possess one (Fig. 4).

As the protease gene families studied here seem likely to contribute substantially to the genomic determinants of phenotypic differences between fungal lineages, the challenge now is to better establish why evolutionary change has taken place in different fungi to determine the cause as well as the function of changes in enzyme profiles. The availability of genome sequences will also facilitate investigations of other paralogous gene families that may give clues as to the potential adaptive significance of lineage-specific expansions in fungi with different lifestyles.

Chapter 3: Variation in gene expression patterns as the insect pathogen Metarhizium anisopliae adapts to different host cuticles or nutrient deprivation in vitro.

Summary

Current molecular and genomic methods are being applied to *Metarhizium anisopliae*, the causative agent of green muscardine disease, because of its importance for biological control of insect pests. It is a very versatile fungus being able to infect a broad range of insects [200 species from > 50 insect families (Samuels *et al.*, 1989)], and is also adapted to life in the root rhizosphere (Hu & St. Leger, 2002). Consistent with its promiscuous nature, an array of ESTs from *M. anisopliae* strain 2575 identified large numbers of genes dedicated to host interaction and countering insect defenses, and regulators for coordinating their implementation (Freimoser *et al.*, 2003). Sequence comparisons and conserved motifs suggest that *ca*. 60% of the ESTs of strain 2575 expressed during growth on cuticle encode secreted enzymes and toxins. Acting collectively the number and diversity of these effectors may be the key to this pathogens ability to infect a wide variety of insects. In contrast, ESTs from the specialized locust pathogen *M. anisopliae* sf. *acridum* strain 324 revealed very few toxins (Freimoser *et al.*, 2003). This relates to life-styles. Strain 2575 kills hosts quickly via toxins and grows saprophytically in the cadaver. In contrast, 324 causes a systemic infection of host tissues before the host dies. This shows that by utilizing ESTs multiple virulence factors and pathways can be viewed

simultaneously, and the different lifestyles that exist in insectfungus interactions can be understood from a broader perspective.

In this report, cDNA microarrays were used for high throughput expression profiling of *how M. anisopliae* strain 2575 responds over a 24 hour period to cuticle from Tobacco hornworm caterpillars (*Manduca sexta*). As a control, we also define the response of *M. anisopliae* to nutrient deprivation. In addition we obtained snap shots of gene expression at 24 hrs to compare and contrast the responses of *M. anisopliae* to gypsy moth caterpillar cuticle (*Lymantria dispar*) and hard (sclerotized) cuticles from a beetle (*Popilla japonica*) or a cockroach (*Blaberus giganteus*).

These studies demonstrated that *M. anisopliae* can rapidly adjusts its genomic expression patterns to adapt to insect cuticle and identified specific responses to different cuticles. Genes specifically induced by cuticle included a plethora of cuticle degrading enzymes, transporters for cuticle degradation products and a subset of transcription factors.

Materials and Methods

Strains and culture conditions

To measure variation in expression of genes during starvation conditions or during adaptation to growth on different insect cuticles we transferred cultures to minimal medium or media containing cuticle after a period of unrestricted growth on a nutrient rich medium. This is an effective and reproducible procedure for obtaining proteins that require release from catabolite repression and/or specific induction by a cuticular component (St. Leger *et al*., 1994). *M. anisopliae sf.*

anisopliae (strain ARSEF 2575) was routinely grown at 27ºC either in liquid (SDB) or on solid (SDA) Sabouraud dextrose medium supplemented with 0.5% yeast extract. For RNA extraction, the fungus was grown for 48 hrs in 50 ml liquid SDB broth. The cultures were then washed with sterile distilled water and 2 g wet weight of the fungal biomass was transferred for up to 24 hrs to 10 ml minimal medium (containing 0.1% (w/v) KH_2PO4, 0.05% (w/v) $MgSO_4$ and 50% (v/v) tap water) supplemented with 1% (w/v) of the following additives: *M. sexta* cuticle (MC, tobacco hornworm), cockroach cuticle (CC, *B. giganteus*), beetle cuticle (BC, *P. japonica*), gypsy moth cuticle (GC, *L. dispar*). Cuticles were prepared as described previously (St. Leger *et al*., 1986b). Alternatively, *M. a. sf. anisopliae* was transferred to 10ml of M. sexta hemolymph (HL) obtained and treated as described (Grundschober *et al*., 1998).

cDNA microarray experiments

All unique ESTs with significant BLAST matches (Freimoser *et al*., 2003) were amplified using T3 and T7 primers and standard PCR protocols. Genes found among the EST sequences of *M. a. sf. acridum* (ARSEF 324) such as chitinases and chitosonase (Freimoser *et al*., 2003), that were absent from the *M. a. sf. anisopliae* (ARSEF2575) EST collection, were amplified from *M. a. sf. anisopliae* genomic DNA with specific primers and included on the array. This resulted in 837 clones that were precipitated and resuspended in 3X SSC (1X SSC: 0.15M sodium chloride, 0.015M sodium citrate, pH 7.0) to give a final DNA concentration between 100 and 300 ng/ml.

Printing, hybridization and scanning of slides was performed with an Affymetrix 417 Arrayer and 418 Scanner (see http://www.umbi.umd.edu/~cab/macore/macorestart.htm for detailed protocols), at the University of Maryland Biotechnology Institute's Microarray Core Facility located at the Center for Biosystems Research. PCR products were spotted in triplicate on poly-lysine coated glass slides with an average spot diameter of 100 μm and spot spacing of 375 μm. Following printing and crosslinking slides were washed with 1% SDS to remove background, treated with blocking solution (0.2 M succinic anhydride, 0.05 M sodium borate prepared in 1-methyl-2-pyrrolidinone) and washed with 95°C water and 95% ethanol. After drying slides were kept in the dark at room temperature.

RNA was extracted as previously described for *M. anisopliae* (Joshi & St. Leger, 1999). For experiments comparing different media, RNA from a culture transferred to SDB was used as the reference sample. For time course experiments mycelia was collected after 1, 2, 4, 8, 12, 18 and 24 h in minimal medium (MM) or from medium containing MC. RNA from the 0 hr time point was used for the reference. Hybridizations were done with Cy3- and Cy5-labeled probes derived from 50-80 mg of total RNA. All hybridizations were repeated at least three times with RNA from independent experiments and with switched labeling for the reference and test RNA samples.

Analysis of microarray data

The images of the scanned slides were analyzed with ScanAlyze (available from the Eisen lab: http://rana.lbl.gov/) and the data obtained from each scanned

slide were normalized using global normalization as performed by J-Express (Dysvik & Jonassen, 2001). All data were log2-transformed and for further analysis the mean (Em) and the standard deviation (SD) of the log transformed expression ratios of the replicates was calculated for all genes. A gene was defined as differently regulated if the expression varied by at least a factor two (1<Em<-1). Expression ratios not fulfilling this requirement (-1<Em<1) were defined as Zero and the same was done for cases where the interval Em ± 1.96xSD (95% confidence interval around the mean value for the three replicate spots) included the value 0. Further analysis of the processed data was performed using J-Express (Dysvik & Jonassen, 2001), EPCLUST (http://ep.ebi.ac.uk/EP/EPCLUST/) and Excel.

Results

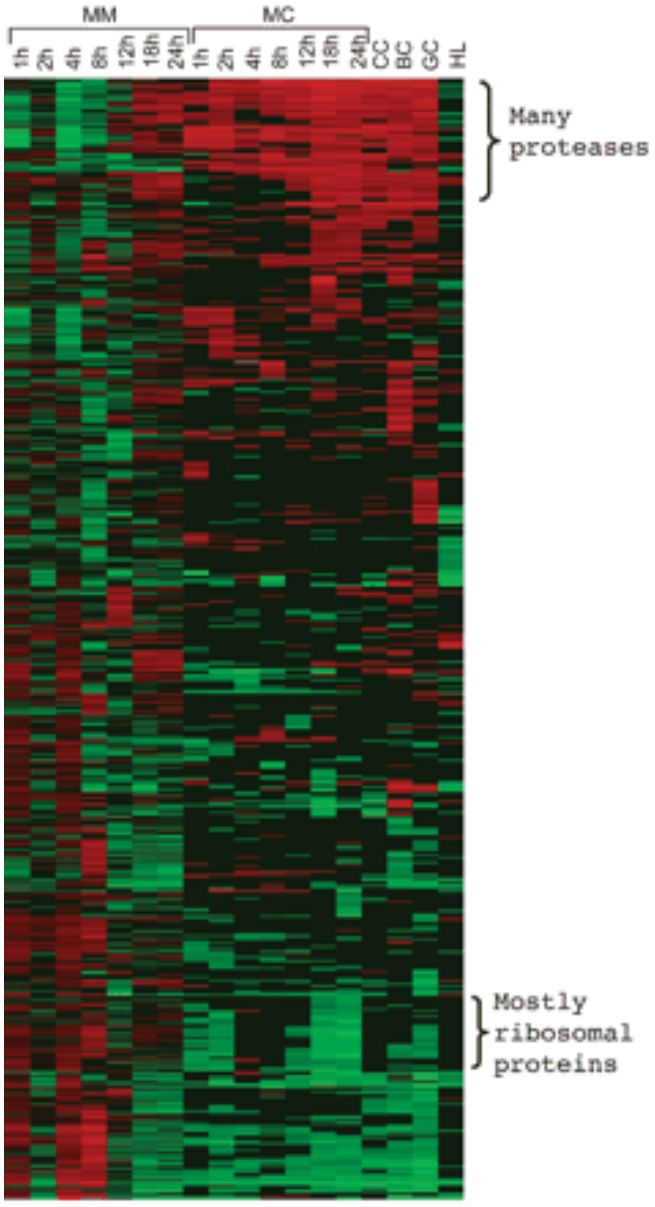
Overall patterns of cuticle-induced gene expression

The libraries we employed to obtain ESTs were made from fungi growing on *M. sexta* cuticle (MC) for 24 h (Freimoser *et al*., 2003). The complete list of ESTs classified into functional groups is available on the internet (http://mic.sgmjournals.org/cgi/content/full/149/1/239/DC1). These ESTs were hybridized with labeled RNA probes isolated from mycelium harvested up to 24 h after the transfer from a nutrient rich medium (SDB) to media containing an insect cuticle or HL. As a control for time course studies with MC, mycelium was challenged by transfer to minimal medium (MM) revealing the response to nutrient deprivation after growth in SDB. The responses to MC and MM were studied in parallel time-course experiments, each with seven time-points (1-24 h) that, together

with the redundant sequence representation in the microarrays, ensured the robustness of the expression profiles.

An overview of the microarray results is presented in Fig 6. They illustrate the rapid changes in expression of some genes in response to MC. Overall, these changes increased in magnitude with time. Thus at 4 h and 18 h post inoculation in MC medium, 88 and 154 genes were up-regulated, respectively. Similarly, 66 genes were downregulated at 8 h and 143 genes were down-regulated at 18 h. During the first hour there was no overlap between genes up-regulated in response to MC and starvation conditions (MM). However, by 18 h and 24 h, 30% of genes were concomitantly upregulated in MC and MM indicating that catabolite repression is involved in regulating at least some cuticle induced genes. A cluster of 41 genes is rapidly activated (< 2 h) by MM but down regulated in response to cuticle (Fig. 1). Only 8 of these genes had homologs of known biological activity in data bases and these included the subtilisin Pr1G and ribosomal proteins. At least in respect of their regulation, nutrient deprivation is perceived as distinct from and even opposite to induction by cuticle.

Fig. 6. Gene expression patterns of *M. anisopliae* in response to starvation conditions (MM), haemolymph (HL) from *Man. sexta*, or cuticles from a beetle (BC), a cockroach (CC) and caterpillars (*Man. sexta*, MC, and *Lymantria dispar*, GC). Mycelia growing on MC or in starvation conditions were assayed in time-course experiments. The 837 cDNA clone set was analysed by hierarchical clustering based on their expression patterns. Genes showing at least twofold regulation, compared with a reference probe from mycelia grown on SDB, are shown in red (upregulated) and green (down-regulated). Colour intensity is directly relative to magnitude of differential expression ratios. Experiments were carried out in triplicate, and representative clusters are shown.



In contrast, the magnitude of expression of most genes up-regulated by nutrient deprivation, including the majority of secreted proteases, was sharply increased by the presence of a cuticle. In addition, a large subset of diverse genes was up-regulated by MC and not MM during the first two hours suggesting that they are specifically involved in adaptation to growth on cuticle (Fig. 7). A broad view of the nature of the adaptations made by *M. anisopliae* following transfer from nutrient rich (SDB) medium to MC was obtained by grouping functionally related genes (Fig. 8). Changes involving up-regulation, measured on the microarray for each functional category during the 24 h of growth on MC were either gradual following the first hour e.g., secreted proteases, or abrupt during the first hour followed by a slow decline, e.g., genes for amino acid/peptide uptake. Down regulated genes included many for protein synthesis machinery, excluding RNA synthesis and processing. Genes encoding ribosomal proteins and translational machinery were coordinately regulated showing an initial decrease, followed by an increase and a decrease. The repression of ribosomal protein genes has been reported in yeasts during multiple stress responses including glucose deprivation (Warner, 1999) and may therefore be a general feature of fungi transferred to a low nutrient medium. Overall, housekeeping genes for cell metabolism, including endocellular proteases, showed stable expression.

Fig. 7. The cluster of *M. anisopliae* genes upregulated within 2 h in medium containing *Man. sexta* cuticle, and which were not upregulated in minimal medium (starvation conditions). Samples were reordered from Fig. 1 according to the time-scale shown across the top, and genes were hierarchically clustered. Gene names and accession numbers are shown to the right of the figure.

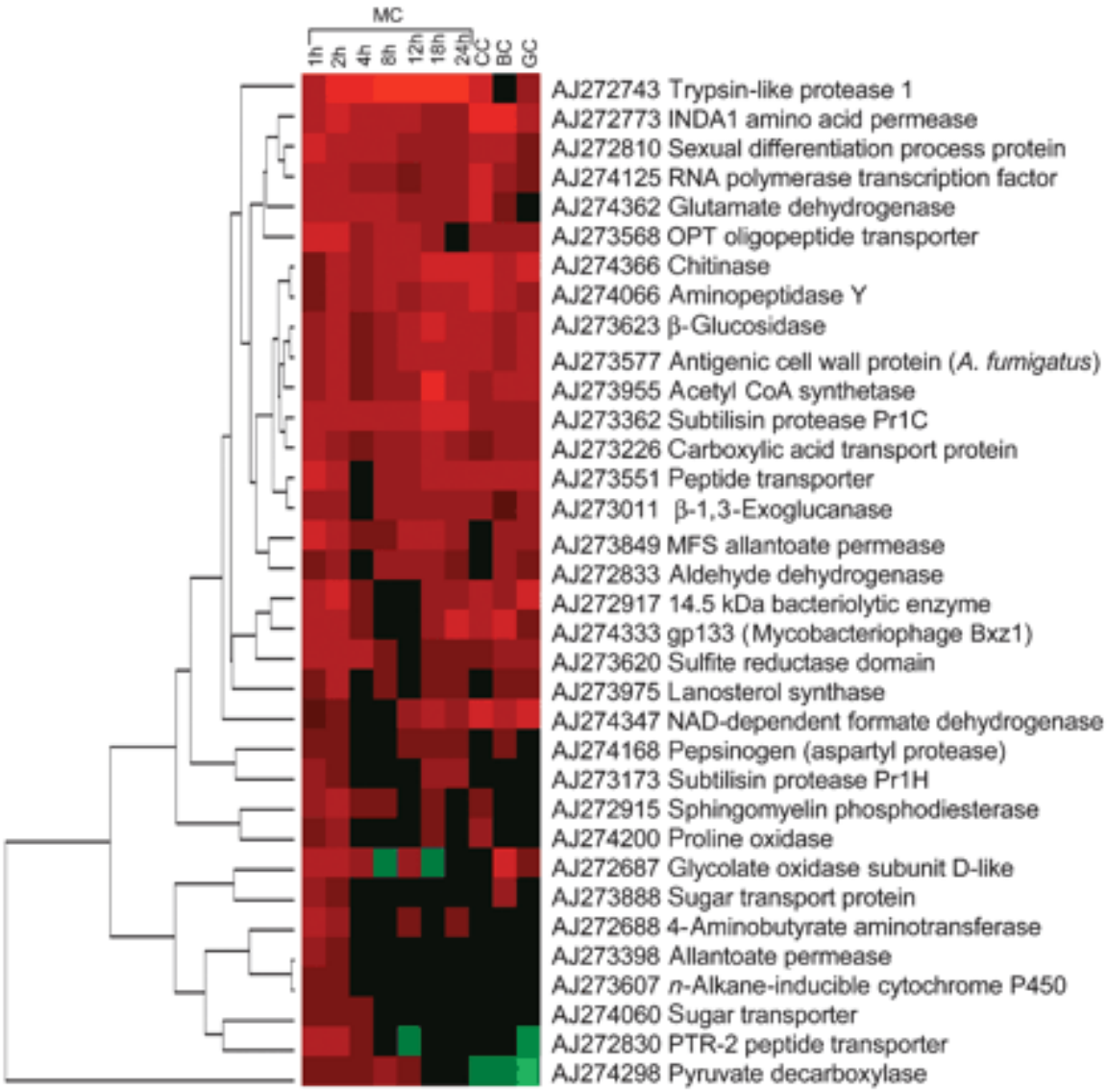
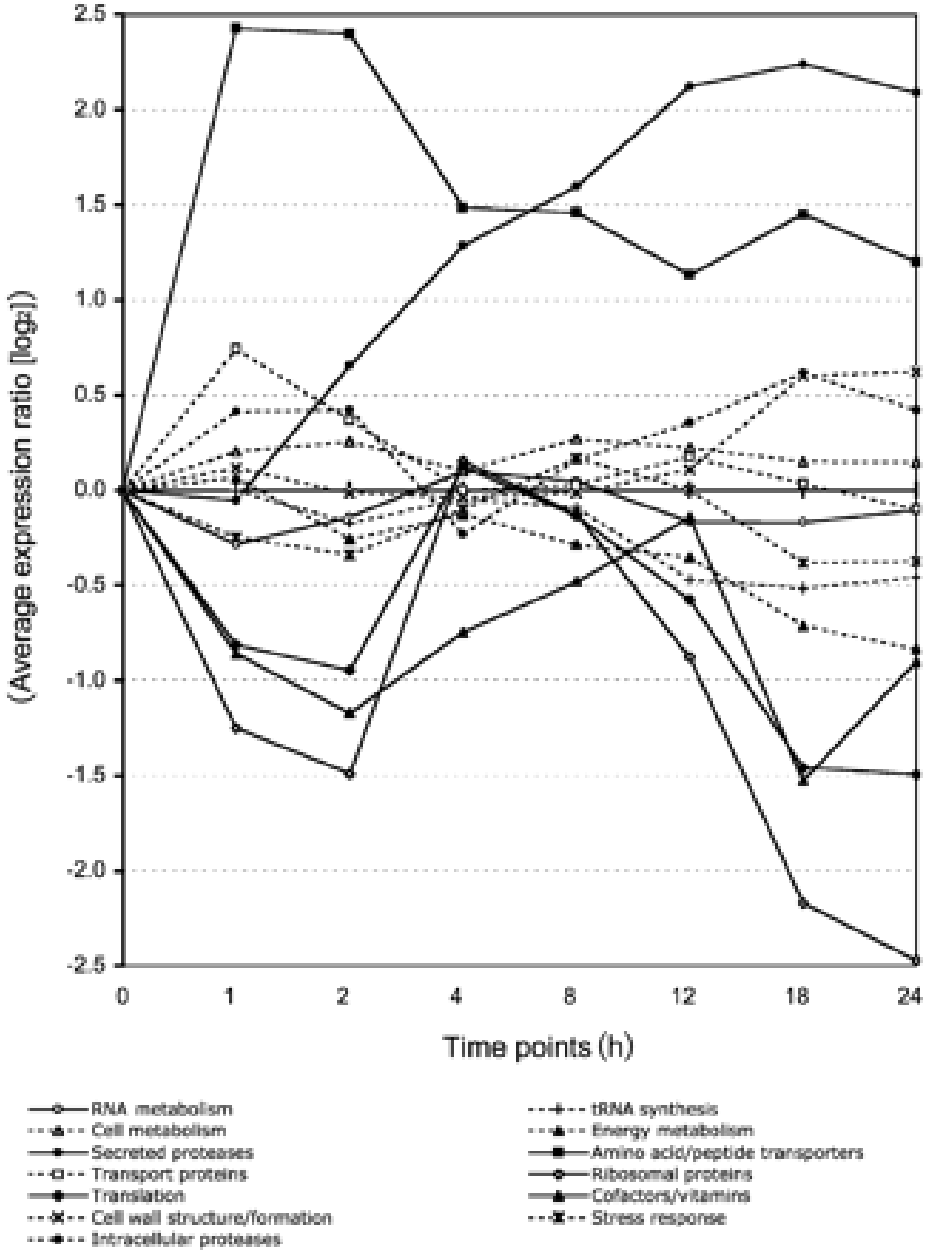
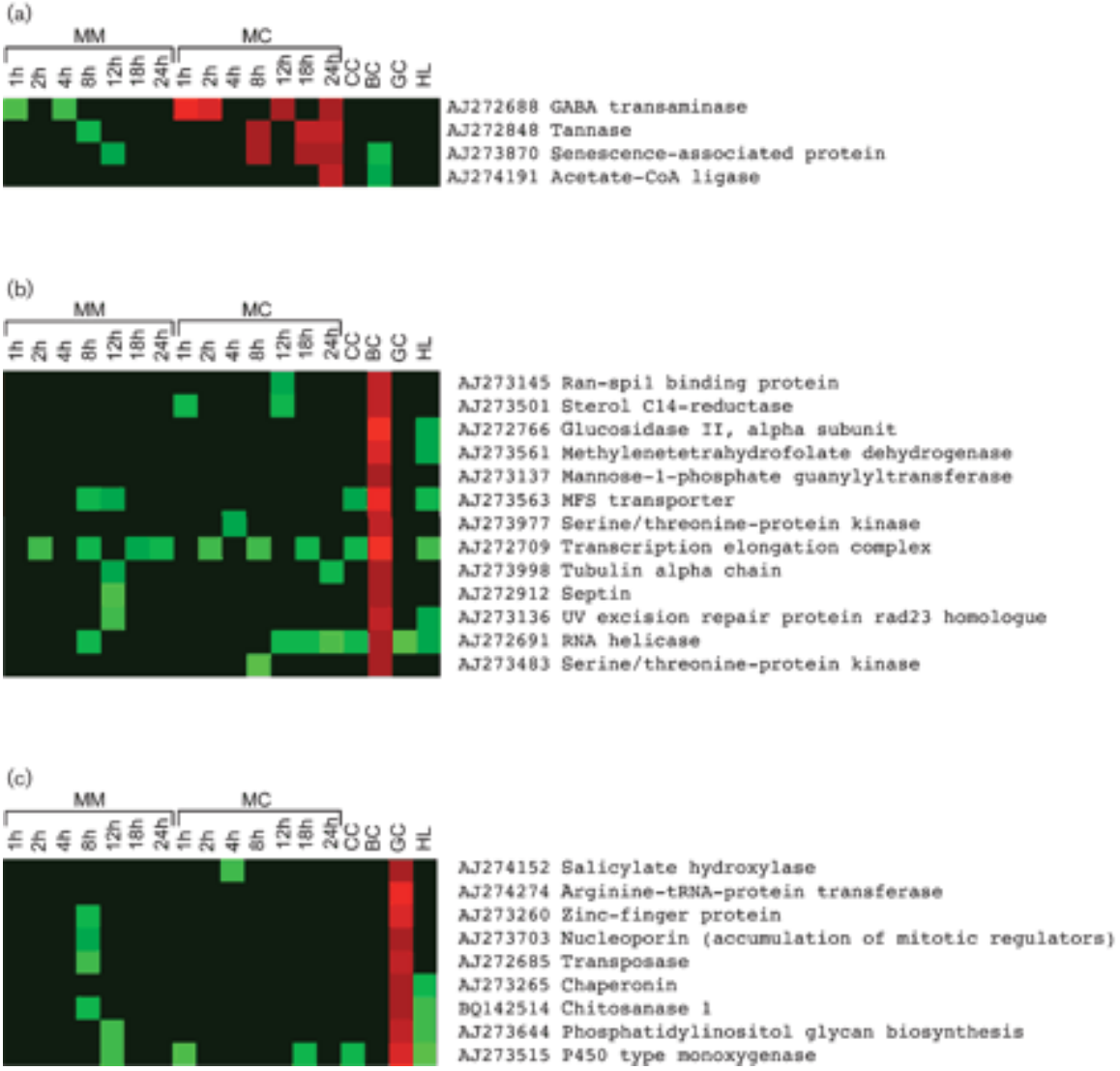


Fig. 8. Regulation of functionally related genes. The curves represent the average induction or repression ratios for all the genes in each indicated group. The total number of genes in each group was as follows: cell metabolism, 71; cofactors/vitamins, 6; energy metabolism, 27; ribosomal proteins, 25; translation, 15; tRNA synthesis, 4; secreted protease, 23; intracellular proteases, 12; transport proteins, 14; amino acid/peptide transporters, 6; cell wall structure/formation, 26; stress response, 26; RNA metabolism, 28.



Based on the time-course experiments under starvation conditions or in minimal medium supplemented with *M. sexta* cuticle (MC), 24 h time point was chosen to obtain a snapshot of gene expression during growth on other insect cuticles. The large scale features of the expression patterns illustrate shared features between the responses to different cuticles as well as starvation conditions indicative of a stereotyped program of gene expression. However, no two expression patterns were identical in terms of the genes effected and the magnitude of expression alteration (Fig. 6). Of the 136 genes up-regulated on MC cuticle at 24 hours, 87 (64%) were similarly regulated on CC, 96 (71%) were similarly regulated on BC and 95 (70%) were similarly regulated on GC. Among these commonly regulated genes 64 were up-regulated on all four cuticles at 24 h. The balance of genes demonstrated specific responses to different cuticles, including up- or down-regulation of genes not observed at any time point on *M. sexta* cuticle (Fig. 9). This implies the pathogen can precisely respond to different conditions. In some cases genes coordinately up-regulated on a particular cuticle were functionally related. Several sequences up-regulated at 24 hours on GC cuticle but not MC, CC or BC (at this time-point) have homologs in yeast involved in integrating nutrient and growth signals with morphogenesis. These include LAS1, a nuclear protein required for cell surface growth and bud formation (Doseff & Arndt, 1995), SLA2 required for morphogenesis and polarization of the membrane cytoskeleton (Holtzman *et al*., 1993), Ecm15p involved in yeast cell wall biogenesis (Goffeau *et al*., 1996) and PIG-L essential in the synthesis of glycosylphosphatidylinositol used as a membrane anchor by cell surface proteins (Watanabe *et al*., 1999).

Fig. 9. Subclusters of genes specifically upregulated on only one of the cuticles (BC, beetle; CC, cockroach; MC, *Man. sexta*; GC, *Lymantria dispar*). Samples were reordered from Fig. 1 according to the time-scale shown across the top, and genes were hierarchically clustered. Gene names and accession numbers are shown to the right of the figure.



Energy Metabolism. While overall expression of several functional categories including cell metabolism, were largely unaltered by transfer to cuticle (Fig. 8), individual ESTs among categories were altered in regulation, possibly indicative of pivotal enzymes involved in metabolic reprogramming. The results of this study compared with experiments performed with *Saccharomyces cerevisiae* (DeRisi *et al.*, 1997) and *Trichoderma reesei* (Chambergo *et al.*, 2002) using very similar nutrient rich and nutrient poor media. In all three species, regulation of many genes that participate in key metabolic processes is not affected by being in sugar rich media such as SDB. However, in *T. reesei*, expression of genes encoding the tricarboxyl acid cycle and mitochondrial proteins favors the oxidation of pyruvate via the TCA cycle rather than its reduction to ethanol by fermentation. In contrast, *S. cerevisiae* preferentially ferments glucose, even in the presence of oxygen. Only when glucose is exhausted do yeast cells use the ethanol as a carbon and energy source for aerobic respiration (the "diauxic shift"). *M. anisopliae* resembled *T. reesei* in that the abundance of transcripts encoding enzymes of the glycolytic pathway and tricarboxyl acid cycle (e.g., isocitrate dehydrogenase, AJ272972) were mostly unaffected upon transfer from a sugar-rich (SDB) to a sugar deficient medium (MM). In yeast these genes are strongly repressed in sugar-rich media. Yeast mitochondrial genes are also subject to strong repression by glucose. However levels of *M. anisopliae* transcripts encoded by the mitochondrial genome (e.g., NADH ubiquinone dehydrogenase, AJ273010) and nuclear genes encoding mitochondrial proteins (e.g., cytochrome C oxidase chain V, AJ272726) were the same or higher in sugar rich media than MM.

These results indicate that like *T. reesei*, but unlike yeast, *M. anisopliae* will respire in the presence of sugar.

However, *M. anisopliae* appears to differ from *T. reesei* in the extent to which aerobic respiration prevails. As in yeast and *T. reesei*, a *M. anisopliae* pyruvate decarboxylase (AJ274332) is up-regulated in the presence of sugar. However, in contrast to these fungi *M. anisopliae* has an additional pyruvate decarboxylase (AJ274298) that is repressed in nutrient rich medium but up-regulated within 1 h on *M. sexta* cuticle (Fig..7). In *S. cerevisiae*, the acetaldehyde formed from pyruvate decarboxylase is reduced to ethanol by alcohol dehydrogenase and is not converted to acetate due to repression of aldehyde dehydrogenase by glucose. Two paralogous genes for aldehyde dehydrogenase have been identified in *T. reesei*, only one of which is repressed by nutrient rich conditions. In contrast, both aldehyde dehydrogenases (AJ272833 and AJ273869) in *M. anisopliae* are down-regulated in SDB as compared to cuticle containing media, suggesting that readily utilized nutrients repress acetate production. It is of interest that AJ272833 is up-regulated earlier on MC than MM (Fig. 7). Two paralogs of acetyl coenzyme A synthetase (AJ273955, AJ274191) were also identified. The AJ273955 transcript is up-regulated early during growth on cuticle and late in MM (Fig. 7), while regulation of AJ274191 is not affected. If both enzymes have comparable specificity, production of acetyl coenzyme A in glucose poor media such as cuticle will increase the entry of acetate, produced via the pyruvate bypass route, into the tricarboxyl acid cycle. Interestingly, *M. anisopliae* also has two paralogous genes for alcohol dehydrogenase. AJ273792 is regulated in a similar fashion to pyruvate decarboxylase AJ274332 (upregulated in

SDB), whereas AJ273547, like pyruvate decarboxylase AJ274298 is repressed in SDB. Thus, *M. anisopliae* has multiple gene families of catabolic enzymes some of which include isoforms that are differentially regulated by sugar. These alternative forms may give *M. anisopliae* the flexibility to shunt any available pyruvate into fermentation or the TCA cycle irrespective of sugar levels.

Amino acid, carbohydrate and lipid metabolism. Genes involved in amino acid catabolism and up-regulated in cuticles included glutaminase A (AJ273512) and NADH-specific glutamate dehydrogenase (AJ274362). Glutamate is the preferred amino acid substrate for *M. anisopliae* (St. Leger *et al.*, 1986a). Otherwise, diverse genes involved in amino acid synthesis were commonly down-regulated in MM and cuticle consistent with reduced availability of raw materials for biosynthesis. Insect cuticle also contains diverse lipids and seven of 13 genes for lipid metabolism were up-regulated on at least one cuticle. Only a cytochrome P450 monooxygenase (AJ274003) was also upregulated during growth in MM. Lipases are the last class of depolymerases to be secreted in insect cuticle (St. Leger *et al.*, 1986b), consistent with which lipase (AJ274124) was up-regulated in late cuticle containing cultures (24 h) only. Enzyme assays also detect a secreted DNAase activity during growth on cuticle (St. Leger *et al.*, 1986b) and DNAase (AJ273950) was up-regulated in cuticle containing media.

Aside protein, the major component of insect cuticle is chitin and predictably therefore chitinases were up-regulated on cuticle. Chitinase AJ274366 was expressed within 1 h on MC, but was not expressed in MM (Fig. 7). Chitosonase was only produced on GC (Fig. 9). As this is coincident with GC-specific expression of genes

involved in morphogenesis is possible that the chitosonase may be involved in modifying cell wall components. However, 5 additional enzymes involved in metabolizing carbohydrates not known to occur in cuticles were also up-regulated in one or more of the cuticle media; formate dehydrogenase (AJ274347) (an activity usually involved in detoxification reactions), 1,2-α -D-mannosidase (AJ273630), β-Dgalactosidase (AJ273808), L-sorbosone dehydrogenase (AJ273834) and β-glucosidase (AJ273623). These could be involved in digesting glycoproteins but were also more weakly up-regulated in starvation conditions consistent with catabolite repression in SDB. Only one of the genes for carbohydrate metabolism (AJ272928) was up-regulated in response to hemolymph, while seven genes were down-regulated (Appendix). Seven carbohydrate metabolizing enzymes were down-regulated on cuticle containing media including a transketolase (AJ274194) and fructose bisphosphate aldolase (AJ273952).

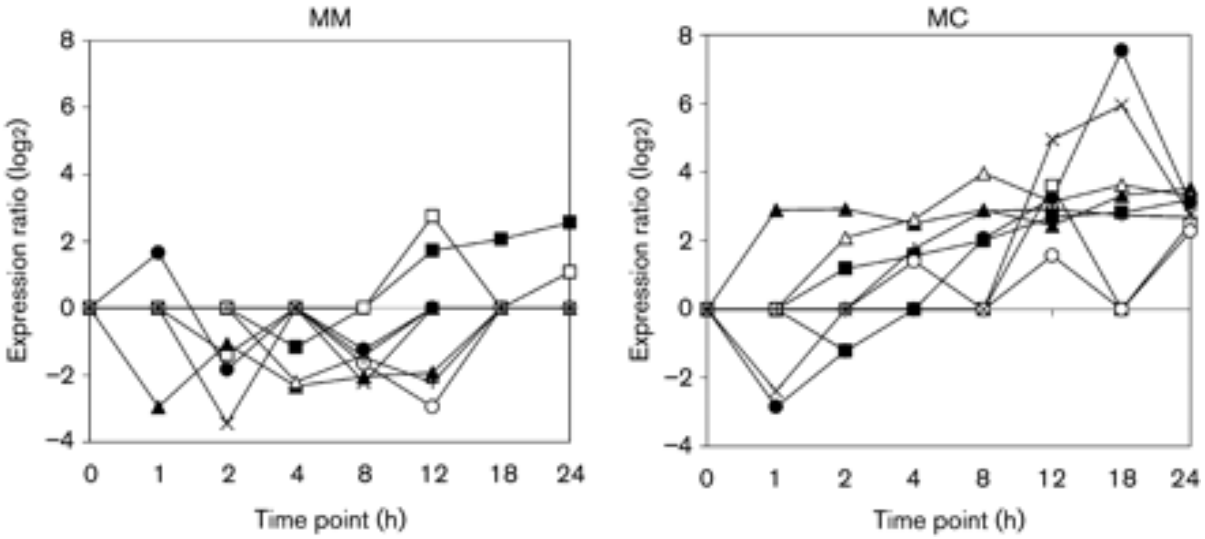
RNA synthesis. Elements required for mRNA synthesis such as RNA polymerase AJ272996 and RNA polymerase transcription factor AJ274125 were up-regulated in cuticle containing media, but not in MM or HL. This presumably adapts the fungus for rapid synthesis of cuticle degrading enzymes.

Transport proteins. The ESTs included two distinct peptide transport systems, one for di-/tripeptides (PTR transporter AJ273551 and PTR-2 transporter AJ272830) and another for tetra-pentapeptides (OPT transporter AJ273568) as well as diverse amino acid transporters (e.g., the INDA1 homolog AJ272773). These all required induction by cuticles and where usually expressed within 1 h (Figs. 7 and 8). In contrast, the PTR transporter in *T. reesei*, is up-regulated by glucose exhaustion alone

(Chambergo *et al.*, 2002) consistent with the *M. anisopliae* transporters having acquired more specialized functions in pathogenicity. Only the *M. anisopliae* oligopeptide transporter OPT2 (AJ273118) was not up regulated in media containing insect cuticle. Regulation of peptide/amino acid transporters was not altered in HL as compared to growth on SDB.

Proteolytic enzymes. It had been shown previously that subtilisin activity is produced in response to nutrient deprivation but that production is enhanced by the addition of cuticle to media (Paterson *et al.*, 1994). Consistent with this, subtilisins *Pr1A* and *Pr1B* were up-regulated on MM as well as to a greater extent on insect cuticles (Fig. 10). Increased induction by cuticle as compared with nutrient deprivation alone suggests that their production is controlled by multiple regulatory systems evoked under different environmental conditions. In contrast, *Pr1C, Pr1D, Pr1E, Pr1F, Pr1I* and *Pr1J* were down-regulated at most time points in MM. Of these, *Pr1C* and *Pr1D* were rapidly up-regulated [*Pr1C* within 1hr of transfer to MC (Fig. 2)], while up-regulation of *Pr1E* and *Pr1K* in MC was delayed by 4 h and 8 h, respectively. *Pr1J* was up-regulated on all the cuticles except BC. *Pr1G* was sharply down regulated in CC. *Pr1F and Pr1I* were up-regulated on MC and on GC. Expression of *Pr1H* was slightly up-regulated by transfer to MC and MM.

Fig. 10. Subtilisin gene expression during growth on cuticle from *Man. sexta* compared with expression during nutrient deprivation. The curves show average expression ratios for different subtilisins in minimal medium (MM) and on Man. sexta cuticle (MC): , Pr1A; , Pr1B; , Pr1I; , Pr1E; +, Pr1J; , Pr1D; , Pr1C; x, Pr1K.



The exo-acting carboxypeptidase AJ274343 was up-regulated after 18h in MM but showed earlier and much stronger up-regulation in all cuticle containing media. Most other categories of exopeptidases (e.g., aminopeptidases AJ273806 and AJ274061) and endo-peptidases including the trypsin (AJ272743), chymotrypsin (AJ273663), metalloprotease (AJ273481) and aspartyl protease (pepsinogen) (AJ274168) were only up-regulated in the presence of cuticle.

Signal transduction. Of the 16 arrayed ESTs encoding proteins involved in transcription, nine (AJ272823, AJ272967, AJ273078, AJ273134, AJ273171, AJ273219, AJ273260, AJ273589 and AJ274235) were up-regulated on at least one cuticle. The positive sulfur transcription regulator homolog (AJ273134) was down regulated in MM and BC, suggestive of particularly low sulfur levels in these media.

Sulfite reductase (AJ273620), but not sulfite oxidase (AJ272866) was up-regulated on cuticles (within 1 hr in MC) but not in MM and HL. In contrast, the pH signaling transcription factor PacC (AJ273219) was up-regulated in cuticles but not in MM or HL. AJ272977 was unique in being up-regulated on HL. In contrast AJ273694 was very strongly down-regulated in HL and strongly up-regulated on the lepidopteran cuticles GC and MC. Although adenylate cyclase (AJ251971) (the enzyme that produces cAMP) and protein kinase A (AF116597) (PKA-the major effecter of cAMP responses) were not up-regulated on cuticle media, a downstream activity, MAP kinase kinase 2 (AJ273356) was upregulated in GC and BC containing media.

Cell wall proteins. Of 30 genes encoding proteins involved in cell structure and function, 18 were up-regulated in at least one cuticle containing medium. Hydrophobins provide another example besides subtilisins where members of a family are differentially regulated, consistent with different functions. Thus, AJ273847 was up-regulated in HL and in MM and down-regulated in cuticle media, while AJ274156 was up-regulated in MM and on sclerotized cuticles (CC and BC), unaltered on lepidopteran cuticles (GC and MC) and down-regulated in HL. This suggests that adaptation to hemolyph may include alterations in cell wall composition. The other cell wall proteins sharply upregulated in both MM and cuticles were AJ273845, a homolog to an antigenic cell wall protein from the human pathogen *Aspergillus fumigatus* and AJ274019 that is very similar to the antifungal glucan 1,3-β -glucosidase from *Trichoderma atroviride* (Donzelli *et al.*, 2001). Clearly, besides cell wall biosynthesis and structure, these proteins may have

additional functions in pathogenicity or in protecting scarce resources from competitors.

Stress response. Several arrayed *M. anisopliae* ESTs are similar to peptide synthases, reductases and other enzymes that take part in the synthesis of fungal toxins such as destruxins, trichothecene and enniatin (Freimoser *et al.*, 2003). This is in agreement with the observation that *M. anisopliae* strain 2575 rapidly kills its host after infection through the action of toxins and subsequently colonizes the insect host by saprobic growth (Samuels *et al.*, 1989). Genes up-regulated in at least one cuticle medium included those encoding a peptide synthase (AJ272930, in BC and GC), a protein involved in sterigmatocystin biosynthesis (AJ273515, in GC), versicolorin B synthase (AJ272697, in CC, GC, MC and HL) and a bacteriolytic enzyme (AJ272917, in CC, BC, GC, early in MC and after 12h in MM).

Discussion

The construction of a *M. anisopliae* cDNA microarray provides many advantages over previous labor-intensive techniques to monitor transcriptional responses to host tissues. For this study it provided a powerful tool with which to examine the influence of culture conditions on the magnitude and spectrum of cuticle-induced gene expression. The analysis presented here has expanded the number of identified *M. anisopliae* genes that respond to cuticle from about 20 (Joshi & St. Leger, 1999; Joshi *et al.*, 1997) to more than 200 genes. Expression patterns of known pathogenicity genes including Pr1A subtilisin, hydrophobin, trypsins, chymotrypsin and carboxypeptidase matched previously published date (Screen & St. Leger, 2000;

St. Leger *et al.*, 1986b; St. Leger *et al.*, 1987; St. Leger *et al.*, 1996; St. Leger *et al.*, 1992). This provides a high level of confidence that the arrays accurately identify differentially expressed clones. For selected genes including subtilisins (Pr1A, Pr1H, Pr1K), a trypsin (Try1) and tubulin, the expression patterns were also verified by quantitative real-time RT-PCR (data not shown). However, as previously observed (Yuen *et al.*, 2002) expression ratios are consistently underestimated by cDNA microarrays as compared to PCR-based methods.

Demonstration of differential regulation of genes encoding cuticle degrading enzymes, cell wall proteins, toxins or toxin producing enzymes on different cuticles, HL and MM suggests that *M. anisopliae* may have the ability to target production of these proteins to different hosts. *M. anisopliae* secretes a greater variety of proteases than reported from any other fungus (Freimoser *et al.*, 2003), some of which have been associated with virulence because they allow rapid physical ingress, nutrient solubilization and the disabling of antimicrobial peptides (St. Leger *et al.*, 1996). The subtilisin cluster provide a good example of *de novo* protein synthesis required for adaptation to growth on cuticle (Fig. 6 and 10), particularly as the differences in regulation of subtilisins imply differences in their function. The proteases such as *Pr1A* produced as part of a general response to nutrient deprivation could also function outside of pathogenesis by scavenging for nutrients during the saprophytic existence. This is consistent with the opportunistic life style of *M. anisopliae.* During early infection processes they could also function in concert with the exopeptidases to provide host degradation products. These may include specialized signals that allow the fungus to "sample" the cuticle and then respond with secretion of the plethora of

cuticle induced proteins. This will include the proteases that require cuticle for induction as they presumably have specialized roles in breaching host barriers. The very early induction of peptide/amino acid transport systems (Figs. 7 and 8) would enhance the ability of the fungus to rapidly and precisely monitor host degradation products.

Of key importance to understanding the mechanisms behind adaptation to cuticles is identification of components of signal transduction that will allow *M. anisopliae* to screen its surroundings to regulate protein synthesis and secretion. PacC-mediated pH signaling is crucial for pathogenicity by the human pathogen *Candida albicans* and the plant pathogen *Fusarium oxysporum* (Caracuel *et al.*, 2003; Davis *et al.*, 2000). Consistent with a crucial role for PacC in *M. anisopliae,* extracellular pH rises during cuticle degradation and acts as a key signal for production of alkaline active enzymes such as subtilisins (St. Leger *et al.*, 1998). Significant by their absence of responsiveness were adenylate cyclase and protein kinase A as transcriptional regulation in response to the cAMP signaling pathways seem central to infection related development in *M. anisopliae* (St. Leger, 1993). Constitutive expression may be a feature of some primary initiators of physiological processes so their importance will not be detected in microarray analyses. A downstream activity, MAP kinase kinase 2 (AJ273356) was up-regulated in GC and BC containing media. This enzyme and other transcription factors may constitute downstream "ground level" components that are immediately concerned with recognizing and responding to specific host features and do not control fungal metabolism as a whole. As such they may be useful for strain improvement purposes.

Microarray technology has made it possible to decipher transcriptional programmes of organisms by studying gene expression *en masse* while assessing individual gene function in a detailed manner (Brown & Botstein, 1999). Thus, knowing when and where a gene is expressed often provides a strong clue as to its function (DeRisi *et al.*, 1997). Almost 50% of the arrayed ESTs up-regulated in cuticle media have undiscovered biological activities and 25% of these are not up-regulated in MM or HL. These have never been recognized to have a role in pathogenicity but are now implicated by co-regulation with known virulence factors. They thus provide an additional rich resource for future research.

Evolutionary theory has long held that the process of adaptation is driven by competition for limited resources. Among heterotrophic microorganisms availability of carbon limits the ability of these organisms to multiply. As a result the machinery of central metabolism is tuned to exploit reduced carbon resources in natural environments where they vary greatly in both form and abundance (Ferea *et al.*, 1999). Comparisons between *M. anisopliae*, *T. reesei* and *S. cerevisiae* suggest that the three fungi will respond differently to environmental changes, presumably reflecting their adaptations to predictable differences in the composition of these environments. The similarities between *M. anisopliae* and *T. reesei* may reflect their close relationship as clavicipitaceous pyrenomycetes. However, the alternatively regulated forms of catabolic enzymes in *M. anisopliae* and *T. reesei* suggest they will differ in how they coordinate regulation of key parts of metabolism such as fermentation at different levels of glucose. This could affect the extent to which aerobic respiration prevails in glucose rich media.

Evidently, gene duplication events and altered patterns of regulation could provide mechanisms for evolution to fine tune ATP-producing pathways as an adaptation of these organisms to their different environments and nutritional requirements. It is tempting to speculate that fermentation may play a more pivitol role in the life of *M. anisopliae* as compared to *T. reesei* to enable it to exploit sugars in the anaerobic environment of the dead host. However, complicating interpretation of these results, ATP-producing pathways can be co-opted to other functions. Thus, some fungal acetyl-coenzyme A synthetases are involved in biosynthesis of secondary metabolites such as penicillin as well as in primary metabolism (Martinez-Blanco *et al.*, 1993). It is axiomatic that as more is learned about the function of each gene, comparative studies on transcriptomes will become an increasingly powerful tool allowing predictive insights into each saprophytes or pathogens behavioral plasticity.

Chapter 4: Field Studies Using a Recombinant Mycoinsecticide (*Metarhizium anisopliae*) overexpressing protease

Introduction

Biocontrol experiments with fungi have often produced inconsistent results, and this has deterred commercial development (St. Leger and Screen 2001). However, many fungi are amenable to genetic modification for purposes of enhancing utility for disease control, insect and plant pest management, or bioremediation. In such cases, genetically engineered fungi may provide environmentally preferred alternatives to current chemical-based control strategies. Much attention has focused on the ascomycete entomopathogen *Metarhizium anisopliae.* It is widely applied abroad, was recently registered for use in the United States and Europe (Butt *et al.* 2001), and offers particular promise as a suppressive agent for many soil insect pests that would otherwise provide a particular challenge to pest control specialists (Milner 1992, Roberts and Hajek 1992). The addition and expression of pesticidal genes in *M. anisopliae* is quite straightforward and was used to genetically engineer a strain that overexpresses toxic proteases and kills insects faster than the wild type dose in laboratory tests (St. Leger *et al*. 1996).

This technology has potential for pest control (St. Leger 2001), but there is an inherent uncertainty about the efficacy, survivability, and environmental risk posed by any introduced or engineered fungus because of our lack of knowledge about the fate of fungal genotypes at the population and ecosystem levels (Bidochka 2001, Hajek 2000). To achieve successful, reproducible, and safe (from the risk management point of view) biological control, we need to be able to study the

ecology of the transformed genotype. After extensive laboratory analysis to test potential risks, including acquisition and evaluation of host range information, an approval (38567-NMP-R) was granted from the Biopesticides and Pollution Prevention Division of the EPA Office of Pesticide Program to conduct a planned release in a field of cabbage plants. The approval constrained the trial to establishing the technology required to monitor the fate of genetically enhanced *M. anisopliae* and to using this technology to determine the potential of engineered strains to establish and disperse over 1-year test period. This was achieved by combining conventional techniques used by soil microbiologists and ecologists with gfp (encoding green fluorescent protein) as a molecular marker. Since root exudates stimulate the growth of bacterial and fungal populations and the rhizosphere is of great importance to plant health and fertility (Whipps 2001), it is a potential refuge for transgenic fungi that could increase their persistence in the environment.

Materials and Method

Fungi and host

The wild-type *M. anisopliae* strain ARSEF1080 was originally isolated from larvae of the cabbage looper (*Trichoplusia ni: Noctuidae*, Lepidoptera) in Florida. Allozyme analysis identified *M. anisopliae* strain 1080 as belonging to genotypic class 14, which is rare in North America (St. Leger *et al.* 1992). The recombinant strain gpd-Pr1-4 contains four copies of the Pr1a subtilisin gene under control of the constitutive gpd promoter from *Aspergillus nidulans* (St. Leger *et al.* 1996).

Transformation

The wild-type strain and gpd-Pr1-4 were transformed with pEGFP-CP (obtained from Don Nuss, Center of Agricultural Biotechnology, University of Maryland, College Park, Md. Plasmid EGFP-CP carries the gene for EGFP1 (a variant of the green fluorescent protein) under control of the glyceraldehyde 3-phosphate dehydrogenase (gpd) promoter from *Cryphonectria parasitica* (Suzuki, Geletka, and Nuss. 2000).

Transformation was performed using a previously established protocol (St. Leger *et al.* 1996) with the modification that inoculated plates were incubated at 28°C for 30 h and transformants visible under a fluorescence microscope were rescued using a glass pipette. Transformants were purified by generating single-spore colonies, and these were subcultured on potato dextrose agar five times to confirm stability. CHEF (clamped homogeneous electric field) gel analysis employing a CHEF-DR-III apparatus (Bio-Rad) was used as described previously (St. Leger *et al.* 1995) to identify transformants carrying the pr1 and egfp1 genes at unlinked locations, i.e., on different chromosomes, so as to allow effective recovery of recombination events. Fluorescent transformant progeny of wild-type (GMa) and gpd-Pr1-4 (GPMa) chosen for the field trial had parent-type growth rate, colony morphology, level of conidial production, and relative virulence as determined by standard laboratory protocols (St. Leger *et al.* 1996).

Fungal release

The field site was located in the University of Maryland Upper Marlboro research farm, Upper Marlboro, Md. It is a frequently cultivated (tilled) site, and the

soil is a Monmouth fine sandy loam. The rectangular 0.2-ha field site was designed to allow for efficient maintenance and the detection of any dispersal of recombinant fungus outside the confines of the plot. The plot consisted of two 0.05-ha fungal application areas, each consisting of seven rows of cabbages separated by a five-row buffer. A barren, plant-free zone surrounded the subplots, and a low-maintenance fallow zone outside the plot was also monitored for marked (recombinant) fungus through the field tests. The cabbage plants (var. Early Flat Dutch) were sprayed on a low-wind day (14 June 2000) with a water-based application containing 0.01% Silwet L77 (Loveland Industries, Greeley, Colo.) at a rate of 10^{13} spores per ha. The ground and the plants in each row were sprayed with a backpack-mounted hydraulic sprayer (18-in. spray band). Application area 1 received GPMa; application area 2 received GMa. Because the purpose of this study was not quantitative, i.e., we were not attempting to compare virulence between GMa and GPMa, a single experimental plot for each was deemed sufficient. Transfer of the fungus by mechanical means was minimized by using a field test design and field test protocol that included the buffer zone and tool and footwear disinfestations.

Collection of soil samples

Before the start of the experiment and daily (first week), weekly (first 2 months), and at monthly intervals thereafter, soil samples were taken at defined depths using a 1-cm soil core sampler from 50, 20, and 15 evenly spaced locations within the application zones, buffer zone, and fallow zone, respectively. Soil samples from the innermost rows of the application areas were taken at 4 to 5 cm from the cabbage tap root as well as alongside the tap root (0 to 1 cm) to check for uneven

distribution and persistence of spores close to the rhizosphere (vicinity of the root). Subsamples of soil were used for dry-weight determination.

Soil samples were stored for up to 3 days at 4°C before the propagules of *M. anisopliae* were quantitated by using Veens semiselective agar medium (Goettel and Inglis 1997.). Soil samples (1 g) were sonicated briefly in 0.05% Tween 80, serial dilutions were made, and 0.1-ml portions were spread on each of two to five plates of selective medium per dilution. The detection limit was less than 20 CFU per g of soil. After range finding experiments, only dilutions near those likely to produce countable numbers of CFU (up to 300 per plate) were plated. The medium was supplemented with hide protein azure to detect constitutive protease production by GPMa (St. Leger *et al*. 1996) and scanned with UV light to distinguish GFP-expressing recombinants from indigenous strains of *Metarhizium spp*. Proc MIXED was used to test for differences in rates of decline of spore titers between GMa and GPMa. Spore count data were transformed to the log scale before analyses. Means were compared using the Student-Newman-Keuls (SNK) method. All analyses were carried out using the SAS software package V8.2. (SAS Institute, Cary NC) (= 0.05).

Indigenous strains of *M. anisopliae* were characterized by allozyme analysis, which allows a large number of strains to be analyzed for recombination events, which we do by assigning a genetic basis to electrophoretic banding patterns (St. Leger *et al.* 1992). Cycloheximide was omitted from the Veens medium to study the abundance and composition of fungal populations other than *M. anisopliae* (total filamentous fungi).

Rhizosphere competence

Samples taken alongside the root by using a 1-cm core borer may contain nonrhizospheric bulk soil that will cause the rhizospheric titer to be underestimated. Therefore, 4 months after planting, eight randomly selected cabbage plants from each application site were cut off above the soil and root samples with adhering soil (rhizosphere samples) were taken. Roots were sectioned into 2-cm segments, and the segments were shaken to collect soil adhering loosely to the roots (outer rhizosphere). To collect soil adhering after shaking but subject to removal by washing (inner rhizosphere), roots were weighed and ultrasonicated (15 s) in sterile water. Outer and inner rhizospheric suspensions were plated onto Veens medium for plate assays. Subsamples of the rhizospheric suspensions were used for dry-weight determinations. To sample rhizoplane *M. anisopliae*, root segments were further washed (10 times), air dried, weighed, and placed on Veens medium.

Monitoring strain stability

Integrative transformants are very stable when grown for long periods in the absence of selection in pure culture under laboratory conditions (Goettel *et al.* 1997, St. Leger *et al.* 1995). However, stability may be different in a complex environment, in which case we reasoned that it would be unlikely for two unlinked markers (Pr1 and GFP) to be lost at once. There should usually be at least one marker remaining to positively distinguish a transformant from a native organism and detect recombination. To determine whether fungi retain the marker elements in their original form, we screened *M. anisopliae* isolates recovered from the application sites for any examples that have lost GFP but retained constitutive

expression of Pr1 or for GFP-expressing strains demonstrating one or more phenotypic characteristics that differ significantly from those exhibited by the input transgenic strains. The growth rate, colony morphology, and level of conidial production were tested as described previously (St. Leger et al. 1996).

Monitoring nontarget arthropods

During the course of the field tests, 50 pit fall traps embedded in the soil in and around the application sites were used to collect nontarget arthropods, particularly carabid beetles (important predators). These were maintained in the laboratory to determine if disease developed, and healthy as well as infected insects were analyzed for the presence of the marked fungus. Insects were placed in petri dishes containing Veens agar medium. Fungal growth over the medium was examined under UV illumination for GFP fluorescence. We anticipated that background levels would be high within the sprayed areas. Consequently, a representative portion (10%) of nontarget insects recovered from these areas were washed briefly in acetone followed by 95% ethanol to remove surface-associated fungal propagules. The individual (identified to species) insects were squashed and placed on Veens agar medium to detect internalized transgenic *M. anisopliae* spores and mycelia. These experiments were designed to determine the extent to which transgenic *M. anisopliae* strains can be recovered from insects (including nonhosts) within an intense deployment area in comparison to the extent found in the surrounding and remote sampling areas, i.e., to determine the potential of insect-mediated dispersal to nontargeted deployment areas.

Result and Discussion

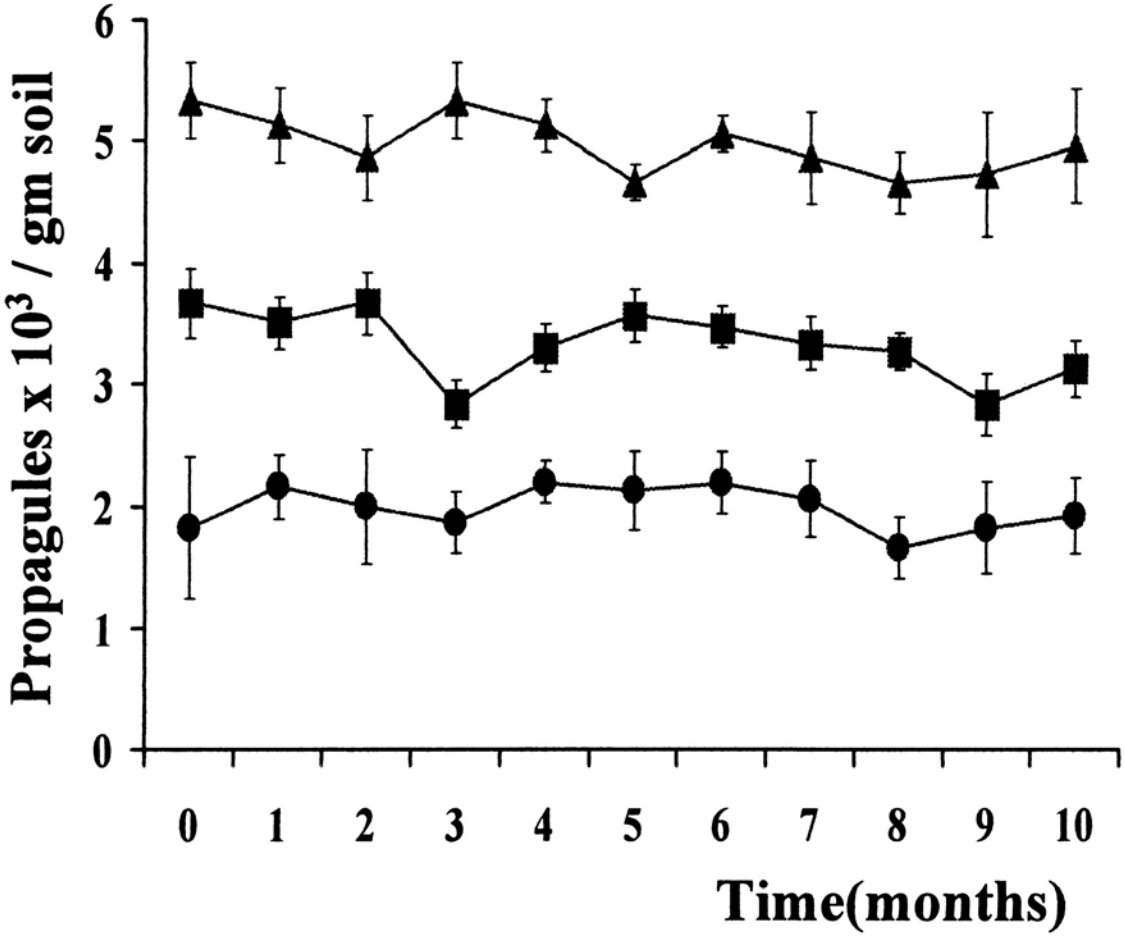
Effects on the indigenous culturable fungal microflora

Before the application, soil samples from the top 3 cm were collected from locations within the application areas, buffer zone, and surrounding fallow zone. Veens medium minus cycloheximide was used to sample total fungal populations. Each fallow location contained at least 106 propagules/g comprising 30 or more fungal species. By contrast, the cultivated buffer and application areas were comparatively impoverished and contained only seven species with a total CFU of less than 104 propagules/g. Veens medium containing cycloheximide grew *M. anisopliae* from only three sites with a mean at these sites of 212 ± 24 propagules/g. Allozyme analysis of 20 colonies picked at random identified two genotypic classes based on electrophoretic phenotypes (St. Leger 1992). Based on assigned genotypes (St. Leger 1992), 14 of the colonies belonged to class 20 (field strain 1). The remaining 6 (field strain 2) differed only at the glutathione reductase locus, which demonstrated a mobility of 121 compared to the 100 shown by class 20. For several months after spraying with transformants, the indigenous strains of *M. anisopliae* were infrequently detected on plates. However, this would be accounted for by the large initial dilution (dilution factor 100) required to obtain countable numbers of GMa and GPMa. Sampling in the spring of the second year of bulk (nonrhizosphere) soil revealed that the original three locations still contained mixed populations of the two indigenous strains. Thus, there is no evidence of a detrimental effect on indigenous populations from introducing GMa and GPMa.

Total filamentous fungal populations in soil 0 to 1 cm from the roots was analyzed using repeated measurements analysis of variance (ANOVA) to determine whether the development of fungal populations is affected by the application of transgenic *M. anisopliae* (interaction between time and treatment). On the Veens medium minus cycloheximide, there was no significant difference in the total rhizospheric fungal populations or its composition in either GPMa- or GMa-treated plants compared to the untreated plants in the buffer zone. In fact, with (Fig. 11) or without (data not shown) application of GMA, the population levels of the three most frequently isolated genera, *Paecilomyces*, *Penicillium*, and *Aureobasidium*, did not change significantly over time.

With the caveat that not all fungi are culturable, these results indicate that there is minimal risk of the engineered fungus displacing naturally occurring fungi. Given the impoverished fungal microflora observed in the cultivated compared to the noncultivated land the impact of introduced microorganisms in general is likely to be minor compared to that of common agricultural practices such as plowing or crop rotation.

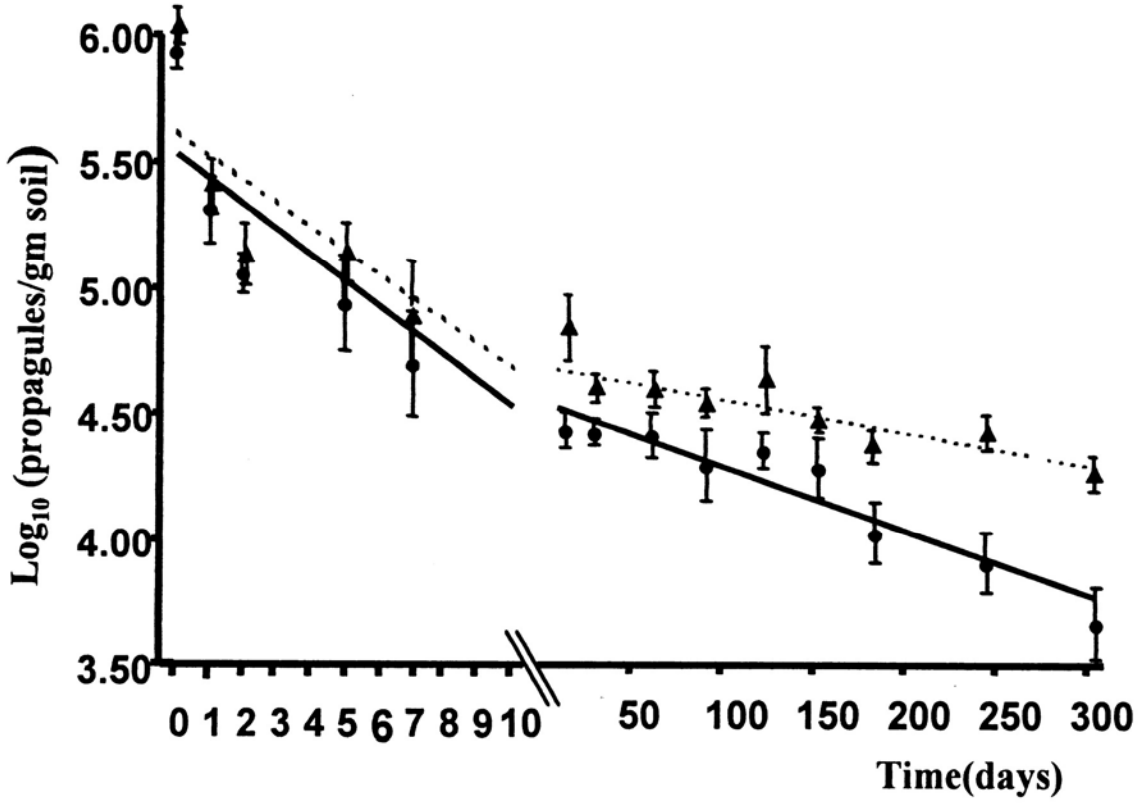
Fig. 11. Mean soil titers of propagules of *Paecilomyces farinosus*, (▲) *Penicillium spp*. (■), *and Aureobasidium pullulans* (●) in the application area treated with *M. anisopliae* GMa. Soil samples were taken within 1 cm from cabbage plant roots at depths of 0 to 2 cm, and fungal propagules were quantified on Veens medium minus cycloheximide. Error bars indicate SD.



Soil persistence monitoring.

At 1 h following application, the differences in the titer of GMa (mean = 2.45×10^5, standard deviation [SD = 6.8×10^4] CFU/g) and GPMa (mean = 2.09×10^5 [SD = 5.2×10^4] CFU/g) were not significant. The data set for subsequent decline in spore numbers in the vicinity of cabbage roots (0 to 1 cm) behaved as piecewise regressions (Fig. 12). Thus, while soil titers of GMa and GPMa declined by 56 and 73%, respectively, in the first week, these strain differences were not significant ($F = 2.9$, $P > 0.05$). Differences between the strains became significant from day 10 ($F = 7.8$, $P < 0.05$). Thus, titers of GMa between 4 months (October) and 10 months (April of the second year) were reduced by 30% from 2.96×10^4, (SD = 1.00×10^3) CFU/g to 2.00×10^4 (SD = 5.29×10^3) CFU/g, respectively. During the same period, titers of GPMa declined by 70% from 1.53×10^4 (SD = 4.16×10^3) to 4.76×10^3 (SD = 1.15×10^3) CFU/g. Reduced fitness and survivability could reasonably be derived from deleterious effects of the additional genetic modifications, compared to the situation for transgenic fungi expressing GFP only. If so, then *M. anisopliae* may not just persist in the soil in a dormant state but characteristics for soil survival may include gene expression that can be interfered with by plasmid integration.

FIG. 12. Changes in the soil titer of GMa (▲) and GPMa (•), over 300 days, within 1 cm of cabbage plant roots at depths of 0 to 2 cm. Spore count data were transformed to the log scale. The lines represent the model outcome of the population decline of GMa (dashed line) and GPMA (solid line). The analysis was conducted using Proc MIXED, SAS. Error bars indicate SD.



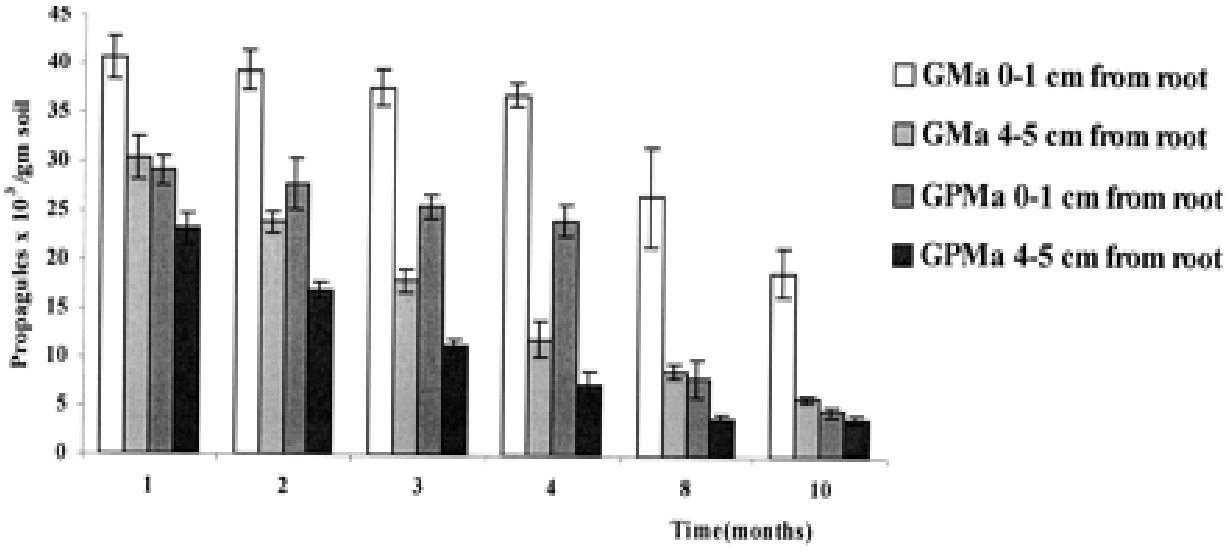
Stability under field conditions

Closely related, therefore, to the issue of population dynamics is the question whether an intensive deployment protocol into complex microbial communities will promote the generation and amplification of altered transgenic *M. anisopliae* strains. Growth rates of colonies isolated from application areas were similar to those of the input transgenic strains. In some cases the isolated colonies differed from their progenitors in producing fluffier, raised colonies and fewer conidia. In each case, the original phenotype returned after three serial propagations on Sabouraud dextrose agar, consistent with experimental and physiological variations rather than genetic differences. During the course of this study, more than 50,000 colonies were examined, and all fungi expressing GFP in application area 1 were also constitutive producers of Pr1 and vice versa. This indicates stability of plasmid DNA in the chromosomes and provides no evidence for recombination. At 10 months after spraying, allozyme analysis was performed on 100 colonies chosen at random from locations also containing field strains 1 and 2. No interisolate variability was detected with each of the eight enzymes possessing electromorphic forms characteristic of genotypic class 14, the class to which strain 1080 belongs (St. Leger *et al.* 1992). In particular, none of the eight loci displayed symmetrical three-banded phenotypes characteristic of heterozygotes or two banded electromorphs, which would be evidence of mixed alternatively homozygous cultures. Spatial distribution of transgenic *M. anisopliae*.

Soil samples were taken at 4 to 5 cm as well as alongside cabbage tap roots (0 to 1 cm) to check for uneven distribution of spores. For GMa, the ratio of fungi

alongside the root to 4 to 5 cm from the root increased from 1.4:1 (1 month postinoculation) to about 3:1 (4 months postinoculation) (Fig. 13). The value remained high after the cabbage plants were killed by frost (8 months), indicating that GMa was persisting on the decaying organic matter. The rhizosphere effect is also apparent in GPMa, although with a more rapid decline in spore titers, reflecting its reduced persistence. Samples from the buffer and fallow zones contained no transgenic *M. anisopliae* or contained insufficient numbers to be detected using dilution plate counts.

FIG. 13. Effect of proximity to cabbage roots on the persistence of *M. anisopliae* under field conditions. Mean soil titers of propagules of *M. anisopliae* GMa and GPMa at depths of 0 to 2 cm are shown. Error bars indicate SD.



Rhizosphere competence

To further analyze rhizosphere competence, samples of soil were taken directly from roots 4 months after application of GMa. A four-fold-larger population of fungal propagules was observed in the inner rhizosphere soil than in the outer rhizosphere soil at the top 2 cm of the root base. This suggests that close proximity to the root and its exudates is involved in the rhizospheric effect. The titer of GMa at the root base (3.1×10^5 [SD = 2.5×10^4] CFU) was close to the original inoculum load. Most other studies using fungi known to be good root colonizers show a decline, perhaps because the initial population added is to large for the carrying capacity of the root . Evidently, soil in the vicinity of plant roots provides a refuge for *M. anisopliae* from factors in the environment that reduce fungal titer.

The colonization of roots by GMa in the outer and inner rhizosphere of roots formed a gradient, with the rhizosphere effect decreasing with increasing depth (Fig. 9). The presence of fungal propagules more than 10 cm from the stem in the inner but not the outer rhizosphere implies some degree of vertical movement along the roots through fungal growth or cracks in the soil or via percolating water. When unwashed root segments were placed on Veens medium, growth of fluorescent fungus was observed from roots up to 10 cm from the stem. In spite of this proximity, a lot of the fungus could be removed from the roots by serial washings (Fig 15). The patchy distribution of fungal colonies on washed roots compared with the total coverage of unwashed roots implies weak adhesion by most propagules.

FIG. 14. Mean number of propagules of *M. anisopliae* GMa at different depths in the outer (■) and inner (□) rhizosphere of cabbage roots, 4 months after application. Error bars indicate SD.

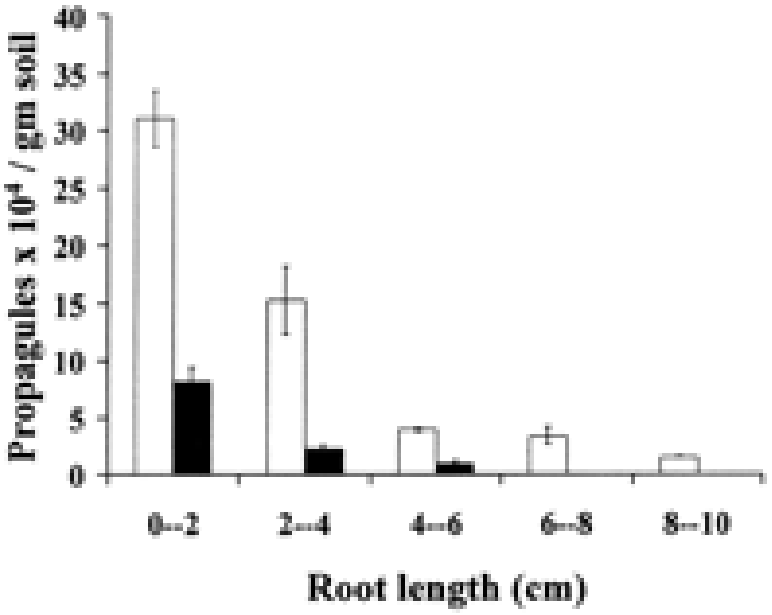
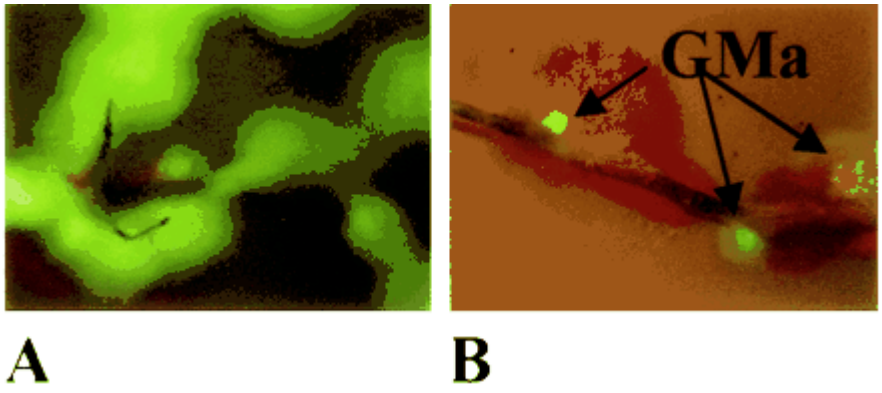


FIG. 15. UV micrographs of 2- to 5-cm-deep cabbage roots from application area 2 (4 months after spraying with GMa) placed on Veens medium for 48 h and showing growth by fluorescent *M. anisopliae*. The roots were shaken free of rhizosphere soil (A) or, in addition, ultrasonicated and subjected to a series of 10 water washes (B), which removed the inner rhizosphere and most but not all fungal propagules (indicated by arrows).



If rhizosphere competence is a general phenomenon among insect pathogens, its impact on plant ecology could be considerable and it possesses implicit coevolutionary implications. In addition, most efforts employing M. anisopliae for biocontrol have ignored habitat preferences and survival outside the host. Evidently, factors associated with soil dwelling may be even more critical in the selection of an isolate than virulence per se (1). However, rhizosphere competence also imposes potential risks since it might increase the difficulty of eliminating a pathogen following unanticipated and deleterious environmental effects.

Monitoring of naturally occurring insects

A natural infestation of *Pieris rapae* occurred on the cabbages. Eight of 20 and 6 of 20 third-instar or older *P. rapae* larvae collected from application areas 1 and 2, respectively, within 10 days of the application died from infections with GPMa and GMa, respectively. A majority of 43 *P. rapae* larvae collected from the test sites 1 month following application died of bacterial infections and parasitoids and further analysis of these was not possible. However, three larvae died from fungal infections and produced spores of GMa. Flea beetles (Alticinae) collected from cabbages up to 1 week after spraying also had spores of GMa and GPMa on their surfaces, but none of 30 beetles maintained for 10 days in the laboratory on cabbage seedlings succumbed to overt fungal infection. More than 3,000 arthropods were collected from pit fall traps in and around the application sites in the summer of the first year. These included four species of carabids (*Amara* and *Stenolophus spp*.), other beetles including predatory rove beetles (*Staphylinidae*), five species of ants, and many types of aphids, springtails, spiders, and mites. About 5% of the arthropods monitored

under laboratory conditions died of a variety of overt fungal infections and 27% died of septicemia or other unidentified causes, but we did not detect external or internalized *M. anisopliae* in healthy, sick, or dead insects. These results suggest the potential of insect-mediated dispersal to nontargeted deployment areas is low.

The results of the field trial do not suggest any safety concerns to using GMa or GPMa that would detract from their being environmentally preferred alternatives to current chemical-based control strategies. However, their survival into the second year is significant, since time increases the possibility of adaptation for increased fitness (Mundt 1995). It cannot be assumed, therefore, that either strain will die out because of current reduced fitness. There was no evidence for phenotypic instability of the introduced fungi, but this might not be expected if genetic changes with clear phenotypes follow the punctuated- equilibrium model of evolution, with long periods of apparent stability punctuated by large infrequent changes. Given that the long-term fitness of a genetically engineered pathogen that persists in nature is difficult to predict, it is all the more essential to establish technologies such as the use of gfp that will permit informed risk assessment through monitoring the fate of marked strains.

Bibliography

Aravinde, L., H. Watanabe, D.J. Lipman and E.V. Koonin. 2000. Lineage specific loss and divergence of functionally linked genes in eukaryotes. *Proc. Natl. Acad. Sci. USA* 97, 11319-11324.

Bagga, S., G. Hu, S.E. Screen and R.J. St. Leger. 2004. Reconstructing the diversification of subtilisins in the pathogenic fungus *Metarhizium anisopliae*. *Gene* 324:159-169.

Berbee M. L. and J. W. Taylor. 2001. Fungal molecular evolution: Gene trees and geologic time. In *The Mycota: A comprehensive treatise on fungi as experimental systems for basic and applied research. Systematics and evolution, Part B*. Edited by D. J. McLaughlin, E. G. McLaughlin, P. A. Lemke, 229-245. New York: Springer

Bidochka, M.J., R.J. St. Leger, A. Stuart and K. Gowanlock. 1999. Nuclear rDNA phylogeny in the fungal genus Verticillium and the relationship to insect and plant virulence, extracellular proteases and carbohydrases. *Microbiology* 145: 955-963.

Braun, E.L., A.L. Halpern, M.A. Nelson and D.O. Natwig. 2000. Large scale comparison of fungal sequence inovation in Neurospora crassa and gene loss in Saccharomyces cerevisiae. *Genome Res.* 10: 416-430.

Brown, P. O. and D. Botstein. 1999. Exploring the new world of the genome with DNA. *microarrays. Nat. Genet.* 21: 33-37.

Brunak, H.N., and G. von Heijne. 1999. Machine learning approaches to the prediction of signal peptides and other protein sorting signals. *Protein Engineering* 12: 3-9.

Bruun, J.A., G. Andrea, E. Jorgen and R. Soren. 1998. Relationships of the insect-pathogenic order entomophthorales (Zygomcota, fugi) based on phylogenetic analyses of nuclear subunit ribosomal DNA sequences (SSU rDNA). *Fungal Gen. Biol.* 24: 325-334.

Caracuel, Z., M. I. Roncero, E. A. Espeso, C. I. Gonzalez-Verdejo, F. I. Garcia-Maceira and A. Di Pietro. 2003. The pH signalling transcription factor PacC controls virulence in the plant pathogen *Fusarium oxysporum*. *Mol Microbiol* 48: 765-779.

Cavalier-Smith, T. 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. *Int. J. Syst. Evol. Microbiol*. 52: 297-354.

Chambergo, F. S., E. D. Bonaccorsi, A. J. Ferreira, A. S.Ramos, J. J.Ferreira, J. Abrahao-Neto, J. P. Farah, and H. El-Dorry. 2002. Elucidation of the metabolic fate of glucose in the filamentous fungus *Trichoderma reesei* using expressed sequence tag (EST) analysis and cDNA microarrays. *J Biol Chem* 277: 13983-13988.

Davis, D., R. B.Wilson and A. P. Mitchell. 2000. RIM101-dependent and-independent pathways govern pH responses in *Candida albicans*. *Mol. Cell. Biol.* 20: 971-978.

DeRisi, J. L., V. R. Iyer and P.O. Brown. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278: 680-686.

Dibb, N.J. and A.J. Newman. 1989. Evidence that introns arose at protosplice sites. *EMBO J*. 8: 2015-2021.

Donzelli, B. G., M. Lorito, F. Scala, and G. E. Harman. 2001. Cloning, sequence andstructure of a gene encoding an antifungal glucan 1,3-beta-glucosidase from *Trichoderma atroviride (T. harzianum). Gene* 277: 199-208.

Doseff, A. I. and K. T. Arndt, 1995. LAS1 is an essential nuclear protein involved in cell morphogenesis and cell surface growth. *Genetics* 141: 857-871.

Duda, T.F.J. and S. R. Palumbi. 1999. Developmental shifts and species selection in gastropods. *Proc. Natl. Acad. Sci. USA*. 96: 10272-10277.

Dysvik, B. and I. Jonassen. 2001. J-Express: exploring gene expression data using *Java. Bioinf.* 17: 369-370.

Eisen, J.A. and C. M. Fraser. 2003. Phylogenomics: intersection of evolution and genomics. *Science* 300: 1706-1707.

Ferea, T. L., D. Botstein, P. O. Brown and R. F. Rosenzweig. 1999. Systematic changes in gene expression pattern following adaptive evolution in yeast. *Proc. Natl.Acad. Sci. USA* 96: 9721-9726.

Freimoser, F. M., S. Screen, S. Bagga, G. Hu and R. J. St. Leger. 2003a. Expressed sequence tag (EST) analysis of two subspecies of *Metarhizium anisopliae* reveals a plethora of secreted proteins with potential activity in insect hosts. *Microbiology* 149: 239-247.

Freimoser, F.M., S. Screen, G. Hu and R.J. St. Leger. 2003b. EST analysis of genes expressed by the zygomycete pathogen Conidiobolus coronatus during growth on insect cuticle. *Microbiology* 149: 1893-1900.

Frugoli, J.A., M.A. PcPeek., T. L. Thomas and C.R. McClung. 1998. Intron loss and gain during evolution of the catalase gene family in angiosperms. *Genetics* 149: 355-365.

Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Hohnston, E. J. Louis, H. W. Memes, Y.

Murakami, P. Philippsen, H. Tettelin, and S. G. Oliver. 1996. Life with 6000 genes. *Science* 274: 563-567.

Graur, D. and W. H. Li. 2000. *Fundamentals of Molecular Evolution, 2nd edition*. Sunderland: Sinauer Associates Inc.

Gruhl, M., W.Y. Kao and G. Bergtrom. 1997. Evolution of orthologous intronless and intron-beraing globin genes in two insect species. *J. Mol. Evol.* 45: 499-508.

Grundschober, A., U. Tuor and M. Aebi. 1998. *In vitro* cultivation and sporulation of *Neozygites parvispora* (Zygomycetes*: Entomophthorales*). *Syst. Appl. Microbiol.* 21: 461-469.

Gunkle, F.A. and H. G. Gassen. 1989. Proteinase K from *Tritirachum album limber*: characterization of the chromosomal gene and expression of the complementary. *Eur. J. Biochem.* 179: 185-194.

Hartung, F., F.R. Blattner and H. Puchta. 2002. Intron gain and loss in the evolution of the conserved eukaryotic recombination machinery. *Nucleic Acids Res.* 30: 5175-5181

Heidelberg, J.F, J.A. Eisen, W.C. Nelson *et al*. (31 co-authors). 2000. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*. 406: 477-483.

Holtzman, D. A., S. Yang, and D. G. Drubin. 1993. Synthetic-lethal interactions identify two novel genes, SLA1 and SLA2, that control membrane cytoskeleton assembly in *Saccharomyces cerevisiae. J. Cell Biol.* 122: 635-644.

Hooper, S.D and O.B. Berg. 2003. On the nature of gene innovation: duplication patterns in microbial genetics. *Mol. Biol. Evol.* 20: 945-954.

Hu, G. and R. J. St. Leger. 2002. Field studies using a recombinant mycoinsecticide (*Metarhizium anisopliae*) reveal that it is rhizosphere competent. *Appl Environ Microbiol* 68: 6383-6387.

Jordan, I.K., K.S. Makarova, J.L. Spouge, Y.I. Wolf and E.V. Koonin. 2001. Lineage-specific gene expansions in bacterial and arcaeal genomes. *Genome Research*.11: 555-565

Joshi, L. and R. J. St. Leger. 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., R. J. St. Leger and D. W. Roberts.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.

Kasuga, T., T.J. White and J.W. Taylor. 2002. Estimation of nucleotide substitution rates in eurotiomycete fungi. *Mol. Biol. Evol.* 19: 2318-2324.

Kondrashov, F.A., I. B. Rogozin, Y. Wolf and E. V. Koonin. 2002. Selection in the evolution of gene duplications. *Genome Biol.* 3: 1-8.

Krylov, D.M., Y. Wolf, I. B. Rogozin and E. V. Koonin. 2003. Gene loss, protein sequence divergence, gene dispensability, expression level, and interactivity are correlated in eukaryotic evolution. *Genome Res.* 13: 2229-2235.

Krzywinski, J and N. J. Besansky. 2002. Frequent intron loss in the white gene: a cautionary tale for phylogeneticists. *Mol. Biol. Evol.* 19: 362-366.

Kumar , S., K. Tamura, I. B. Jakobsen and M. Nei. 2001. MEGA2: Molecular Evolutionary Genetics Analysis software., Arizona State University, Tempe, Arizona, USA.

Martinez-Blanco, H., M. Orejas, A. Reglero, J. M. Luengo and M. A. Penalva.1993. Characterisation of the gene encoding acetyl-CoA synthetase in *Penicillium chrysogenum*: conservation of intron position in plectomycetes. *Gene* 130: 265-270.

Murphy J. M. and J. D. Walton. 1996 Three extracellular proteases from *Cochliobolus carbonum*: Cloning and targeted disruption of ALP1. *Molecular Plant-Microbe Interactions.* 9: 290-297.

Nei, M. and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol*. 3, 418-426.

Nikoh, N., and T. Fukatsu. 2001. Evolutionary dynamics of multiple group I introns in nuclear ribosomal RNA genes of endoparasitic fungi of the genus Cordyceps. *Mol. Biol. Evol.* 18: 1631-1642.

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

Rawlings, N.D. and A.J. Barrett. 1995. Families of serine peptidases. *Methods Enzymol.* 244: 19-60.

Ross, J., H. Jiang., M. Kanost and Y. Wang. 2003. Serine proteases and their homologs in the *Drosophila melanogaster* genome: an initial analysis of sequence conservation and phylogenetic relationships *Gene* 304: 117-131.

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

Screen, S. E. and R. J. St. Leger. 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.

Sharp, P.M., and W.H. Li. 1987. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 15: 1281-1295.

Siezen, R.J. and J. A.M. Leunissen. 1997. Subtilases: the superfamily of subtilisn-like serine proteases. *Protein Science* 6: 501-523.

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

St. Leger, R.J., R. M. Cooper, and A. K. Charnley. 1987. Distribution of chymoelastases and trypsin-like enzymes in five species of entomopathogenic deuteromycetes. *Arch. Biochem. Biophys.* 258: 123-131.

St. Leger, R. J., B. May, L. L. Allee, D. C. Frank, R. C. Staples and D. W. Roberts. 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. and S. E. Screen. 2001. Prospects for strain improvement of fungal pathogens of insects and weeds. In *Fungal Biocontrol Agents: Progress, Problems and Potential,* edited by T.M. Butt, C. Jackson and N. Morgan, 219–238. Wallingford : CAB International

St. Leger, R. J. 1993. Biology and mechanisms of insect cuticle invasion by deuteromycete fungal pathogens. In *Parasites and Pathogens of Insects*. edited by N. C. Beckage, S. N. Thompson and B. A. Federici, 211-229. London : Academic Press.

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

St. Leger, R. J., Joshi, L. & Roberts, D. 1998. Ambient pH is a major determinant in the expression of cuticle-degrading enzymes and hydrophobin by *Metarhizium anisopliae. Appl Environ Microbiol* 64, 709-713.

St. Leger, R. J., Joshi, L., Bidochka, M. J., Rizzo, N. W. & Roberts, D. W. 1996. 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. and 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.

St. Leger, R.J., Joshi, L., Bidochka, M.J., Rizzo, N.W., and Roberts, D.W. 1996. Construction of an improved mycoinsecticide over-expressing a toxic protease. *Proc. Natl. Acad. Sci. U. S. A*. 93: 6349–6354

Tajima, F. 1993. Simple methods for testing molecular clock hypothesis. *Genetics* 135: 599-607

Takahashi, K. and M. Nei. 2000. Efficiencies of fast algorithms of phylogenetic inference under the criteria of maximum parsimony, minimum evolution, and maximum likelihood when a large number of sequences are used. *Mol. Biol. Evol.* 17: 1251-1258.

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

Tilburn, J., S. Sarker, D.A. Widdick, E.A. Espreso, M. Orejas, J. Mungroo, M.A. Penalva, H.N. Arst jr. 1995. The *aspergillus* PacC zinc finger transcription finger mediates regulation of both acid- and alkaline-expressed genes by ambient pH. *EMBO J.* 14: 779-790.

Walton, J.D. 1996. Host selective toxins: Agents of compatibility. *Plant Cell* 8: 1723-1733.

Warner, J. R. 1999. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci* 24: 437-440.

Watanabe, R., K. Ohishi, Y. Maeda, N. Nakamura and T. Kinoshita. 1999. Mammalian PIG-L and its yeast homologue Gpi12p are N-

acetylglucosaminylphosphatidylinositol de-N-acetylases essential in glycosylphosphatidylinositol biosynthesis. *Biochem. J.* 339: 185-192.

Yuen, T., E. Wurmbach, R. L. Pfeffer, B. J. Ebersole and S. C. Sealfon. 2002. Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. *Nucleic Acids Res.* 30: e48.

Zhang, J., Y.P. H.F. Rosenberg and M. Nei. 1998. Positive Darwinian selection after gene duplication in primate ribonuclease genes. *Proc. Natl. Acad. Sci. USA* 95: 3708-3713.