

Abstract

JEONG, JUN SEOP. Genomic studies of the rice blast fungus, *Magnaporthe grisea*.

(Under the direction of Dr. Ralph A. Dean.)

Magnaporthe grisea, the rice blast pathogen, is a model plant pathogenic fungus. Extensive genetic research and rich structural genomic studies on this fungus have opened the door into the era of functional genomics. Expression-based genomic technologies represent the core of functional genomics. cDNA libraries are central part of genomic studies for gene discovery and functional studies such as protein expression. We developed a novel and efficient method for the generation of normalized full-length enriched cDNA libraries by combining SMART with recombinational cloning. Sequence analysis indicated that the normalized cDNA library contained novel transcripts not found previously by large scale EST sequencing. Proteins that are secreted or presented on the outside of the fungal cell wall likely represent primary determinants of pathogenesis. Microarray analysis of microsomal-associated and free cytosolic RNAs was performed to identify secreted and membrane-associated proteins. The data suggested that some proteins may be translocated to the secretory pathway post-translationally. Also, we observed differential enrichment of transcripts under different growth conditions, indicating that protein secretion might be a highly regulated process. Signal pathways play a central role in fungal-host interactions. 14-3-3 is a major regulator of signal transduction pathways; however, its function has not been clearly defined. Deletion of MgFTT1, a 14-3-3 homolog, in *M. grisea* resulted in reduced pathogenicity. Mutants also

exhibited alteration in cell wall properties and extracellular carbohydrate hydrolase activities, consistent with the observation that deletion strains were defective in the ability to grow *in planta*. To improve the rate of homologous recombination in *M. grisea*, we developed a split-marker strategy for gene replacement. Split marker was at least 10 fold more efficient for deletion of a phytotoxic snodprot1 homolog, MgSPH1, than using whole construct-mediated transformation. The deletion mutant showed no obvious defect in appressorium formation. However, the mutants exhibited reduced pathogenicity, which at least in part was due to a defect in growth *in planta*. Application of purified MgSPH1p to wounded leaf tissue did not cause any apparent phytotoxic effects as has been shown for other snodprot1 homologs. In conclusion, we developed a variety of genomics tools and strategies to begin to functionally investigate the molecular basis of fungal pathogenicity. This work also contributes significantly to the functional annotation of the rice blast fungus genome.

GENOMIC STUDIES OF THE RICE BLAST FUNGUS, *Magnaporthe grisea*

by

Jun Seop Jeong

A dissertation submitted to the Graduate Faculty of

North Carolina State University

In partial fulfillment of the

Requirements for the degree of

Doctor of philosophy

Functional Genomics

Raleigh, NC

Spring 2006

Approved by:

Ralph A. Dean

Gary A. Payne

David F. Ritchie

Jeffrey L. Thorne

Dedication

This work is dedicated to my family. The support and trust from my parents are beyond description. Also, JiSu Bang has been a great support and love during this journey. My daughter, Ena M. Jeong, was born in 2005. My parents, my wife, and Ena taught me how precious life could be. Thank you all for your love and support throughout my journey.

Biography

Jun Seop Jeong was born on June 11, 1971 in Suwon, South Korea. He attended Seoul National University from 1990 to 1994 for Bachelor of Science in Agriculture. After graduation, he served in army for 2 and half years as an officer. After the completion of military service, he joined a research group led by Dr. Yong-Hwan Lee as a master student. Soon after the graduation with Master of Science in Agriculture in 1998, he moved onto Genomics to join Dr. Ralph A. Dean's research group. Since then, he has been working on a variety of genomic aspects of the rice blast fungus.

Acknowledgement

I would like to express my sincere appreciation to my major advisor Dr. Dean. With deep insight, he has led me through my graduate research. I appreciate the advice from my committee members. The work in Dr. Payne's lab inspired me to work on the characterization of 14-3-3 in the rice blast fungus. I would like to thank Dr. Ritchie for his willingness to serve on my committee. Also, I appreciate the knowledge and insight from Dr. Thorne. Most of my computational and statistical knowledge came from him. A lot of people in FGL have come and gone during my graduate study. They helped and encouraged me to cope with the ups and downs on my research. I wish to especially thank Tom for being such a good friend. I would like to thank Nichole for helping me with the microarray work. I would like to thank Dr. Eisenharber for analyzing GPI anchored protein identification, Doug for coordinating the process, and Dr. Deising for the gift of antibody that I used for my research. I would like to thank Dr. Lee for having given me advice on the path that I took and that I will continue to take.

Table of Contents

List of Tables	viii
List of Figures	ix
Chapter One: Literature Review	1
Overview of the rice blast fungus	2
The infection process	2
Features of the <i>M. grisea</i> genome	4
Expressed sequence tag (EST) as a reagent for the functional genomics	5
Functional analysis of genes in the rice blast fungus	7
Signal transduction pathways for appressorium formation	10
Secreted proteins and pathogenesis in the rice blast pathosystem	12
Current approaches in identifying secreted proteins	15
Application of microarray for molecular localization and functional genomics	19
Conclusion	21
References	23
Chapter Two: A novel strategy for cDNA library generation using SMART (Switching mechanism At 5' end of RNA template) and recombinational cloning systems	43
Abstract	44
Introduction	45
Materials and methods	48
Total RNA preparation	48
cDNA generation and cloning	48
cDNA library normalization	49
Sequence analysis	50
Results	50
cDNA library construction strategy	50
cDNA library normalization	51
Cloning	51
Sequencing	52
Discussion	53
References	56
Figure legends	65
Chapter Three: Genome-wide microarray analysis of secretome in the rice blast fungus, <i>Magnaporthe grisea</i>	72
Abstract	73
Introduction	74
Materials and methods	78

Strains and culture.....	78
RNA sample preparation.....	79
Microarray analysis and validation.....	80
Immunofluorescence microscopy.....	81
Results.....	81
Analysis of microarray data.....	82
Computational prediction of proteins with signal peptide and membrane-bound proteins.....	84
Factors affecting MBP enrichment.....	85
Prediction of GPI-anchored proteins.....	88
Transposable element transcripts are associated with the microsomal fraction.....	89
Discussion.....	90
References.....	95
Figure legends.....	115
Chapter Four: 14-3-3 homolog, MgFTT1, in <i>Magnaporthe grisea</i> affects carbon metabolism and is required for growth <i>in planta</i>	128
Abstract.....	129
Introduction.....	130
Materials and methods.....	133
Fungal strains and culture condition.....	133
Identification of MgFTT in the genome.....	134
Generation of Knockout mutants.....	134
Microscopic studies.....	136
Appressorium assays.....	136
Enzymatic assay.....	137
Pathogenicity assay.....	137
Results.....	138
Identification of MgFTT1.....	138
Expression analysis of MgFTT1.....	139
Generation of knockout mutants.....	139
Deletion of 14-3-3 affects the expression of hydrolytic enzymes.....	140
Deletion of MgFTT1 does not affect the ability to form appressorium.....	141
MgFTT1 is required for growth <i>in planta</i>	142
14-3-3 GFP fusion protein localizes to the nucleus.....	143
Discussion.....	144
References.....	154
Figure legends.....	166
Chapter Five: Gene disruption using split-marker in the rice blast fungus, <i>Magnaporthe grisea</i>	178
Abstract.....	179
Introduction.....	180
Materials and methods.....	182

Culture, gene identification, and sequence analysis	182
Gene deletion by split marker	183
Appressorium and pathogenicity assays	185
Expression studies.....	185
Results.....	186
Sequence characterization of MgSPH1	186
Split-marker was efficient in <i>in vivo</i> gene deletion	188
MgSPH1 is dispensable for appressorium formation	189
MgSPH1 is required for pathogenicity	190
MgSPH1 recombinant protein does not appear to be phytotoxic	190
Expression of MgSPH1 is sensitive to nutrient starvation.....	191
Discussion.....	192
References.....	196
Figure legends.....	203
General Conclusion.....	213
References.....	217
Appendix.....	220
Secretome display	221
Objective	221
Materials and Methods.....	221
Results and discussion	223
References.....	225
Figure legends.....	226

List of Tables

Chapter Two: A novel strategy for cDNA library generation using SMART (Switching mechanism At 5' end of RNA template) and recombinational cloning systems

Table 2. 1 BLAST analysis of non-normalized and normalized cDNA libraries.....	62
Table 2. 2 BLAST analysis of cDNA sequences.....	63

Chapter Three: Genome-wide microarray analysis of secretome in the rice blast fungus, *Magnaporthe grisea*

Table 3. 1 Primers used for RT-PCR analysis.....	106
Table 3. 2 List of Top 25 transcripts enriched in MBP and FP in MM48 condition.....	107
Table 3. 3 List of Top 25 transcripts enriched in MBP and FP in APP condition.....	109
Table 3. 4 Statistical significance for signal ratio for signal peptide and transmembrane domains	111
Table 3. 5 List of GPI-anchored proteins in the rice blast fungus genome	112

Chapter Five: Gene disruption using split-marker in the rice blast fungus, *Magnaporthe grisea*

Table 5. 1 Screening of transformants for homologous recombination.....	202
---	-----

List of Figures

Chapter Two: A novel strategy for cDNA library generation using SMART (Switching mechanism At 5' end of RNA template) and recombinational cloning systems

Figure 2. 1 cDNA synthesis scheme	67
Figure 2. 2 Comparison of cDNA generation.....	68
Figure 2. 3 cDNA library normalization.....	69
Figure 2. 4 cDNA filter hybridization of normalized cDNA library	70
Figure 2. 5 The size distribution of the non-normalized and normalized cDNA libraries.....	71

Chapter Three: Genome-wide microarray analysis of secretome in the rice blast fungus, *Magnaporthe grisea*

Figure 3. 1 Sample preparation for microarray analysis.....	119
Figure 3. 2 Partitioning of the genes in ergosterol biosynthesis pathway	120
Figure 3. 3 The distribution of MBP:FP signal ratio for transcripts encoding proteins with signal peptide	121
Figure 3. 4 RT-PCR analysis of microarray result	122
Figure 3. 5 Comparison of hydropathy plot of signal peptides.....	124
Figure 3. 6 Translational activities of select <i>M. grisea</i> transcripts	125
Figure 3. 7 Differential staining of chitin and chitosan during infection.....	126
Figure 3. 8 The distribution of MBP:FP signal ratio for MAGGY components (gag and pol) grown for 48 hrs in minimal medium	127

Chapter Four: 14-3-3 homolog, MgFTT1, in *Magnaporthe grisea* affects carbon metabolism and is required for growth *in planta*

Figure 4. 1 Phylogenetic tree using fungal 14-3-3 RefSeq sequences retrieved from GenBank	170
Figure 4. 2 Expression pattern of MgFTT1 in different growth stages and conditions..	171
Figure 4. 3 Targeted deletion of MgFTT1	172
Figure 4. 4 Enzymatic activities in culture filtrates	173
Figure 4. 5 Deletion of MgFTT1 affects cell wall properties	174
Figure 4. 6 Pathogenicity assays of wild type and mutant strains	175
Figure 4. 7 The localization of MgFTT1::GFP fusion protein	176
Figure 4. 8 Over-expression of MgFTT1::GFP affects appressorial development	177

Chapter Five: Gene disruption using split-marker in the rice blast fungus, *Magnaporthe grisea*

Figure 5. 1 The strategy for the deletion of MgSPH1 by split-marker gene replacement system	206
---	-----

Figure 5. 2 Sequence and phylogenetic analyses of MgSPH1 homologs.....	207
Figure 5. 3 Kyte-Doolittle hydropathy plots of hydrophobins and MgSPH1 homologs	208
Figure 5. 4 Deletion of MgSPH1 was confirmed by PCR.....	209
Figure 5. 5 MgSPH1 is required for fungal pathogenicity.....	210
Figure 5. 6 Purification of recombinant MgSPH1 protein.....	211
Figure 5. 7 Expression of MgSPH1 is nutrient-sensitive.....	212

Appendix: Secretome display

Figure A. 1 Expression of Aga2p as a C-terminal fusion.....	228
Figure A. 2 Expression of MPG1 test constructs and FACS analysis.....	229
Figure A. 3 Enrichment of cells expressing proteins with signal peptide by FACS and MACS	230

Chapter One

Literature Review

Overview of the rice blast fungus

The rice blast pathogen, *Magnaporthe grisea*, is a filamentous ascomycetous fungus. Rice blast is one of the most destructive diseases in rice production worldwide. Disease management relies principally on the breeding of resistant cultivars. However, resistance is transient due to the emergence of new races of the rice blast fungus. About half of the world's population relies on rice as a major food (Goff, 1999). More than 90% of the world's rice is grown and consumed in Asia, and considerably more rice production is anticipated due to the rapid population growth in this part of the world (Khush, 2005). Due to its economic importance as a staple food, the fungus has a long and rich history of research. The isolation of fertile strains greatly facilitated establishment of genetic research on the fungus. Also, development of genetic transformation system has enabled functional analysis of pathogenicity related genes (Leung et al., 1990; Valent and Chumley, 1991). More recently, a complete genome sequence has become available. Consequently, rice blast has emerged as a leading model system to study fungal-host interactions.

The infection process

The infection process of the rice blast fungus begins with a sequence of morphogenetic events that culminates in the formation of an appressorium. The formation of functional appressorium is a prerequisite for successful colonization of host. Extensive research on the appressorium formation and overall infection process has been prolific, revealing pathways that are important for the appressorium formation in the rice blast fungus (Dean, 1997). The appressorium is formed from the short germ tube emerging from conidium,

which can be distinguished from hyphopodia that emanate from hyphae (Emmett and Parbery, 1975). The external cues for appressorium formation in plant pathogenic fungi may vary. In the rust fungus, *Uromyces appendiculatus*, for example, appressorial development is known to be triggered by topology of the contact surface which mimics the ridge of the guard cell of the host (Hoch et al., 1987). The fungus gains entry into host through stomatal openings. This mechanism of thigmotropic recognition may be specialized for the rust fungus. In many fungi, including *M. grisea*, hydrophobicity of surface has been shown to be an important factor in appressorium formation, again, which mimics host surfaces. However, hydrophobicity does not appear to be essential for certain fungi such as *Erysiphe graminis* (Nicholson, 1996). Along with physical properties of the surface, substances from host leaves have been implicated in chemical cues for appressorium formation. This was exemplified by ethylene, a plant hormone for ripening, in *Colletotrichum* spp. and cutin monomers in *M. grisea* (Flaishman and Kolattukudy, 1994; Gilbert et al., 1996). It might be that fungal pathogens have specialized infection stimuli adapted from their niches. Recently, the rice blast fungus was shown to be able to adapt the strategy of root infecting fungi forming hyphopodia when infecting underground parts of the host plant (Sesma and Osbourn, 2004). Thus, fungal pathogens may have adapted to the ecological niches they occupy to infect host plants. This adaptation can be sophisticated to meet their survival as in rust fungi while other generalist pathogens such as *Colletotrichum* species might be less meticulous in their recognition of environmental stimuli that result in infection structure formation.

Features of the *M. grisea* genome

The genome of the rice blast fungus is about 40MB. The rice blast fungus represents the first plant pathogenic fungus to be sequenced (Dean et al., 2005). There are estimated to be over 13,000 genes in the haploid genome. The genome sequence has revealed the presence of a large number of repetitive elements such as transposons and retrotransposons in the genome. In rice production, durable resistance to the rice blast is hampered by the emergence of new races. Transposable elements are one of the causal agents of genetic variability (Fudal et al., 2005; Kang et al., 2001; Orbach et al., 2000). Rice infecting strains are characterized with abundance of repetitive elements. This might indicate that the presence of transposable elements in the genome may confer a genetic flexibility to the plant pathogen. Another pronounced feature of the genome is the presence of genes that might be important for sensing and communicating with the environment, which was manifested by expansion of CFEM domain proteins and putative G-protein coupled receptors (GPCRs) (Kulkarni et al., 2005). Often species specific gene expansion is associated with the adaptation to the ecological niches. Compared to its close relative *Neurospora crassa* (Galagan, 2003), the rice blast fungus is known to have a larger repertoire of putative G-protein coupled receptors. Among those putative GPCR, specific expansion of CFEM domain proteins are evident. Some CFEM domain proteins have 7 transmembrane domains consistent with GPCRs. To support their role in pathogenesis, pth11 was shown to be required for pathogenicity (DeZwaan et al., 1999), and a CFEM domain protein, ACI1, was shown to bind MAC1, adenylyl cyclase in *M. grisea* (Kulkarni and Dean, 2004). MAC1 is also required for cAMP synthesis and pathogenesis. The genome of two rice varieties have been published, recently (Goff,

2002; Yu, 2002). The sequence of both rice and rice blast makes it the first plant host and fungus combination available to investigate the interaction at the genomic level. Comparison of gene content between the rice blast fungus and saprophytic *N. crassa* revealed a number of interesting aspects. There are less proposed secreted proteins in the *N. crassa* compared to the *M. grisea*, suggesting pathogen-specific genomic adaptation. This is further supported by the presence of many consensus carbohydrate substrate-binding domains. Analysis of *M. grisea* genome indicates an abundance of many pathogenesis-related genes such as polyketide synthases, non-ribosomal peptide synthases, enzymes required for terpenoid metabolism, ABC transporters, and MFS transporters. These represent common themes in pathogenesis by plant pathogenic fungi.

Expressed sequence tag (EST) as a reagent for the functional genomics

cDNA library generation and accompanying EST sequencing is a central part of genomic studies (Adams et al., 1991; Rudd, 2003). Since cDNA is generated from RNA, the information contained in cDNA libraries is essential for the understanding of dynamics of the transcriptome. Comparative EST analyses can provide insight into physiology of host-pathogen interaction (Kahmann and Basse, 2001; Skinner et al., 2001), and help prioritize analysis of genes using a spectrum of other functional genomics tools (Sweigard and Ebbole, 2001). In higher eukaryotes, the transcriptional unit is unusually monocistronic. Consequently, the full length cDNA library generation and EST sequencing is an important step toward defining gene structure for *in silico* studies such as gene prediction, transcriptional start site mapping, and defining promoters (Das et al., 2001; The FANTOM Consortium et al., 2005). However, construction of full length

cDNA libraries have been hampered by the presence of premature termination of Reverse Transcriptase (RT) reaction during first strand synthesis (Carninci et al., 1998). Moreover, the inherent redundancy of abundant transcripts makes EST sequencing not cost-efficient toward for finding the full complement of expressed genes (Bonaldo et al., 1996). Although, engineered versions of RT, for example RHase H-negative strain, provides greater processivity to reach 5' end of transcripts, it is also not straightforward to determine if the resulting cDNA is full length. To this end, enzymatic manipulation of transcripts before the first strand synthesis or capture of full length cDNA after first strand synthesis has been attempted to construct cDNA library (Carninci et al., 2000). In addition, affinity-based enrichment of full-length transcripts before cDNA generation has been attempted (Edery et al., 1995). Another approach, oligo-capping, exploited the tobacco acid phosphatase to tag eukaryote-specific cap structure with a known oligonucleotide (Maruyama and Sugano, 1994). However, these methods often require demanding techniques, require a large amount of starting material, and need processing of RNA samples prior to cDNA synthesis. A simple procedure to synthesize full-length enriched cDNA library construction method was devised using intrinsic terminal transferase activity of reverse transcriptase, termed SMART (Switching Mechanism At 5'end of RNA Template) (Zhu et al., 2001). The protocol requires a minimal amount of starting material, and employs a template switching mechanism which is enhanced by the 5' cap structure (Schmidt and Mueller, 1999). Cloning is one of the critical steps in the construction of cDNA libraries. The DNA ligase-based ligation step is subject to variation. And, the selection of recombinant cloning is typically done using blue/white selection using α -complementation. Recently, recombinational cloning strategy was

introduced (Ohara and Temple, 2001). The cloning is mediated by lambda-based integration and excision apparatus. Recombinant clones are selected on the basis of negative selection. The cloning system was shown to have several advantages; it is more robust and provides a higher level of recombination compared to conventional ligation, a larger number of recombinants due to the selection on the replacement of a toxic gene in the cloning vector, less size bias, and ease of transfer of inserts to an array of utility vectors without gain or loss of nucleotides. Since the initial cloning step requires small recombination sequences it can be easily constructed using a PCR-based approach. Thus, the combination of SMART and the recombinational system would provide a powerful and dependable strategy to construct deep coverage full length enriched cDNA libraries.

Functional analysis of genes in the rice blast fungus

Now that the field of structural genomics has become relatively mature, functional genomics has surfaced with the aim of analyzing data from large scale sequencing, and is rapidly evolving with advent of new technologies and paradigms. Early genomic studies in the rice blast fungus produced a solid frame work such as cDNA library generation, EST sequencing, large insert cosmid and BAC library construction. Targeted gene knockout system is well established in the rice blast fungus. Due to the fact that many filamentous fungi spend most of their life cycle in the haploid form, the effect of gene deletion is often direct if the targeted gene is a single copy. Equipped with the feasibility of gene targeting in the filamentous fungi compared to other higher eukaryotes, targeted gene deletion has become a primary functional analysis tool. However, the rate of homologous recombination and efficiency of DNA delivery into the cell remains far from

optimal. Homologous recombination is often more efficient with longer (>1kb) matching sequences. The cloning of long DNA sequence into a targeting construct can be a major bottleneck. To help address this problem, the TAG-KO system was developed (Hamer et al., 2001). The system employs transposon mutagenesis of large insert DNA in *E. coli* which can then be used for homologous recombination in *M. grisea*. This technique may be advantageous in saturation level mutagenesis of chromosomal regions of interest. Also, long flanking sequences increase the rate of homologous recombination. However, the technique is not well suited for high throughput analysis and presents a number of technical challenges. To overcome relatively low transformation efficiency, *Agrobacterium tumefaciens*-mediated transformation (ATMT) was developed for filamentous fungi, and was shown to dramatically increase the number of transformants recovered (de Groot et al., 1998). ATMT has been used in the rice blast fungus for insertional mutagenesis and detailed conditions for transformation have been developed (Rho et al., 2001).

To increase the frequency of legitimate recombination and reduce the effort needed to identify gene knockout transformants, two novel strategies have been developed. The deletion of Ku70 and Ku80 which function in homologous end-joining of double-stranded DNA breaks was shown to increase the rate of homologous recombination dramatically in *N. crassa* (Ninomiya et al., 2004). This results in a significant reduction in the number of transformants that need to be screened to identify gene knockouts. The strategy is promising for large scale mutagenesis studies. A second strategy employs the use of a suicide gene such as HSV tk for positive-negative selection in combination with a pro-drug such as Ganciclovir (Capecchi, 1989; Mansour et al., 1988). Since the current

gene deletion strategy depends on double homologous recombination flanking target region, ectopic and/or single integration event is likely to retain the suicide gene expression cassette, enabling both negative and positive selection for sensitivity to pro-drug and antibiotic resistance, respectively. The strategy was successfully employed in *Fusarium oxysporum* and *M. grisea* (Khang et al., 2005).

The systems described above require cloning of targeting constructs although alternatives have been developed. Budding yeast exhibits a high rate of homologous recombination even with a short stretch of sequence homology. This facilitates the creation of PCR-mediated deletion amplicons, obviating the cloning step (Wach et al., 1994). Adaptamer-mediated PCR has proven to be a powerful tool for facilitating efficient gene targeting in yeast (Reid et al., 2002). The adaptamer-PCR strategy allows cloning-free generation of deletion constructs, incorporation of relatively long flanking sequences, and incorporation of various selectable markers. However, the ability of Taq polymerase to amplify long sequences is limited. Another strategy, referred to a split-marker, may serve as an efficient tool for gene deletion in filamentous fungi (Fairhead et al., 1996). Split-marker relies on homologous recombination within the selectable marker, which most likely occurs when the flanking sequences are in proximity, and presumably when both DNA fragments have undergone homologous recombination with genomic DNA. As a result, it is expected that most transformants exhibit knockout of the target gene. Moreover, the sequences to be amplified by PCR are smaller compared to the whole PCR amplicon by adaptamer-PCR, enabling more efficient generation of deletion constructs.

Signal transduction pathways for appressorium formation

As fungi grow, they are continuously monitoring their environment. Outcome of this communication may result in the elaboration of varying morphology and developmental structures. Often, such developmental processes are linked to nutritional sources, resulting in, for example, filamentous growth or formation of sexual structures. In budding yeast, nitrogen starvation with a fermentable carbon supply results in the formation of hyphae (Gimeno et al., 1992). Also, depletion of glucose serves as a cue for haploid invasive growth in yeast (Cullen and Sprague, Jr., 2000). This process is closely linked to intracellular signal transduction mediated by cAMP (Pan and Heitman, 1999). Also, pheromone-responsive MAPK pathway is involved in polarized growth (Roberts and Mosch, 1997). Such a dimorphic switch is not uncommon in pathogenic fungi. Many animal and plant pathogenic fungi often exhibit a dimorphic switch, and this morphological transition is often linked to pathogenicity. cAMP is known to be a universal secondary messenger. cAMP induces morphogenesis in a variety of fungi. The central effector of cAMP is protein kinase A. The role of cAMP in pathogenic fungi has been firmly established. Mitogen activated protein kinase (MAPK) pathway is another crucial component of morphological development in pathogenic fungi.

cAMP induces appressorium formation in the rice blast fungus on a non-inductive hydrophilic surface (Lee and Dean, 1993). PKA catalytic subunit deletion mutants exhibit abnormal appressoria and reduced pathogenicity (Mitchell and Dean, 1995) (Xu et al., 1997). Extensive research suggests that PKA and MAPK kinase pathways may interact as well as act independently. Filamentous growth and invasive growth are the hallmark phenotypes associated with PKA and MAPK pathways in yeast. Characterization of PKA

catalytic subunits and STE12 in budding yeast suggested independent roles for PKA in unipolar budding and MAPK in cell elongation with both affecting invasive growth (Pan and Heitman, 1999). In the rice blast fungus, it was also shown that the two pathways play a distinct role in nutrient mobilization (Thines et al., 2000). Recent studies to dissect PMK1-related MAPK pathway suggest that divergent input and output of the pathway might exist even with apparent conservation in the modular structure of the MAPK pathway (Li et al., 2004; Park et al., 2002; Zhao et al., 2005). Interestingly, MST12, a STE12 homolog in the rice blast fungus, appears to have PKA phosphorylation sites, implying interplay between cAMP/PKA and MAPK pathways.

14-3-3 is an interesting molecule that has a potential role in the interaction of MAPK and PKA (Dumaz and Marais, 2005). 14-3-3 was first identified as a protein that is abundant in brain tissue. Subsequently, it has been found to play an indispensable role in many signal transduction pathways, and is a universal protein in eukaryotes. Putative roles of the 14-3-3 protein includes allosteric enzyme modulator, phosphoprotein binding activity, scaffolding of proteins, regulation of subcellular localization, cell cycle control, and exocytosis. So far, functional studies in fungi are limited. In budding yeast, deletion of both copies of 14-3-3 is lethal. However, in a certain genetic background, deletion of both copies of 14-3-3 was shown to be viable (Roberts and Mosch, 1997). The deletion strain (*bmh1/bmh2*) showed pleiotropic phenotypes such as sensitivity to genotoxic substance, and pseudohyphal growth, suggesting a role in cell cycle progression and polarized growth. Also, the double deletion strains showed an altered ability to grow on alternative carbon source and sensitivity to osmotic stress. However, individual deletion of either *bmh1* or *bmh2* did not exhibit dramatic sensitivity to those conditions,

suggesting functional redundancy. Interestingly, BMH1p and BMH2p were shown to bind an upstream element of pheromone-responsive MAP kinase component, STE20p, a p21-activated kinase. In *A. nidulans*, overexpression affected conidiation and polarized growth (Kraus et al., 2002). The expression of a truncated construct in *Trichoderma reesei* rescued a secretion defect phenotype and showed elevated invertase activity in budding yeast (Vasara et al., 2002). The multi-faceted role of the protein is manifested through modulation of enzyme activity, modulation of transcription factor activity (Beck and Hall, 1999; Igarashi et al., 2001), and regulation of nucleocytoplasmic shuttling of cdc25 (Graves et al., 2001; Lopez-Girona et al., 1999). Importantly, 14-3-3 proteins modulate MAPK kinase pathway through PAK (Roberts and Mosch, 1997), raf (Irie et al., 1994), MAPKKK (Adams et al., 2002; Fanger et al., 1998), and possibly downstream of a GTPase Rac1 (Hurtado and Rachubinski, 2002). Genetic interaction of yeast 14-3-3, BMH1 and BMH2, with PKA (Gelperin et al., 1995) suggest that 14-3-3 protein is a participant in actin organization through MAP kinase and cAMP/PKA components (Roth et al., 1999) (Eby et al., 1998; Holly and Blumer, 1999). However, a role for 14-3-3 in fungal phytopathogenic fungi has not been determined.

Secreted proteins and pathogenesis in the rice blast pathosystem

Secreted proteins play a pivotal role in plant-microbe interactions. Although this class of proteins includes a broad range of functionality such as regulatory roles, enzymatic activity, and structural roles, the function of secreted proteins in the biotic interaction is ambiguous. Secreted proteins have been of central interest in plant pathology. Early studies focused on hydrolytic enzymes which degrade host components during infection

and colonization. Such hydrolytic enzymes serve as virulence factors rather than pathogenicity factors (Schafer, 1994). In the rice blast fungus, a cutinase gene, CUT1, was disrupted without affecting pathogenicity (Sweigard et al., 1992). Also, deletion of two xylanases, representing about 80% of xylanase activity in stationary phase of growth, did not affect infection of the fungus (Wu et al., 1997), and the study revealed additional xylanase activities. This substantiates the view that cell wall degrading enzymes play ancillary roles to “basic compatibility factors” (Walton, 1994). Thus, the importance of secreted cell wall degrading enzymes in fungal pathogens is not clearly defined. However, recent characterization of SNF1 kinase homolog in *Cochliobolus carbonum* and *Fusarium oxysporum* suggests that overall cell wall degrading enzymes may affect pathogenicity in plant pathogenic fungi (Ospina-Giraldo et al., 2003; Tonukari et al., 2000). Interestingly, deletion of a MAP kinase Gpmk1, a PMK1 homolog, in *F. graminearum*, caused delayed induction of a subset of secreted hydrolytic enzymes that might be important for the infection process (Jenczmionka and Schafer, 2005). However, the cause and consequences of the observation is not clear.

In contrast to enzymes that degrade the host cell wall, other studies revealed that the disease relationship between the pathogen and host plant is determined by dominant genetic factors, termed avirulence genes in the pathogen and resistance genes in the host. The concept was elaborated by Keen, proposing that the avirulence-R-gene interaction is mediated by a ligand-receptor mechanism (Keen, 1990). However, direct physical interaction between the avr gene product and the R gene product has been rarely found. Thus, this model was subsequently modified to yield the guard model (Dangl and Jones, 2001). The guard model proposes that the avr gene product forms a complex with its

virulence target. The virulence target is guarded by an R protein. Thus, when an avr product enters the host cell, and interacts with its virulence target, this complex is recognized by the R protein, which activates resistance responses. Using extensive map based cloning, an avirulence gene, avr-Pita was cloned in *M. grisea* (Orbach et al., 2000). As a rare case, physical interaction of Avr-Pita and Pita gene was shown using yeast two hybrid, and a mutation harbored in the susceptible host cultivar prevented the physical interaction (Bryan et al., 2000; Jia et al., 2000). It is not known how the Avr-Pita gains access to the rice cytoplasm for physical interaction. However, interesting insight may be gained from comparative studies in the malarial parasite and Oomycetes. Research in the malarial pathogen, *Plasmodium falciparum*, revealed molecular machinery that delivers parasitic proteins into host cells, termed the host cell targeting (HCT) signal (Hiller et al., 2004). This process is essential for remodeling the host cell, protection from the host immune response, and establishment of parasitism. Recent studies of Oomycetes avirulence gene products unveiled the presence of similar motifs in protein sequences of avirulence gene products and the possibility of a similar mechanism Oomycete-plant interaction (Armstrong et al., 2005; Birch et al., 2006). This is consistent with the cytoplasmic localization of many R-gene products hypothesis where by avirulence gene products are delivered into host cytoplasm. Thus, these findings shed light on a possible existence of evolutionarily conserved infection mechanism in both plant and animal pathogens. However, it is not known if true fungal avirulence gene products are delivered into host cytoplasm in a similar manner. This will be a fascinating future area of research. Even though no universal conserved structural motifs have been found in avirulence gene products, it has been suggested that a surveillance system in host plants may detect

function such as enzymatic activity, rather than structural motifs (Shao et al., 2003). This is further supported by the fact that many avirulence gene products are proteases. Moreover, protein turnover is an important theme with regard to the activation of plant defense systems (Austin et al., 2002). This might lead to targeted and more specific research for screening of avirulence gene activity.

Recently, a targeted approach to the identification of secreted protein during pathogenic interaction was focused on haustorially expressed secreted proteins from a rust pathogen, revealing enrichment of potential avirulence genes (Catanzariti et al., 2006). Secreted proteins represent a broad functional category ranging non-specific cell wall degrading enzymes to specific avirulence gene products. Thus, it is important to identify secreted proteins in a target specific manner from a pathogen to better understand both mechanism of pathogenesis and defense.

Current approaches in identifying secreted proteins

Most secreted proteins harbor a signal peptide at the N-terminus. Signal peptides are short, and have 3 distinct regions with regards to the physical properties of the amino acids. Heavy dependence has been placed on the use of signal prediction software. SignalP is one of the most widely used software tools (Bendtsen et al., 2004). Some potential problems of using such software with annotated genomes include i) inaccurate genome annotation, ii) different translation start sites *in vivo*, iii) exportation of proteins out of the cell using different mechanisms. The current status of gene predictions may be accurate with regard to defining regions with gene coding potential. However, delineating borders such as exon-intron borders and 5' and 3' end of genes remains challenging. This

may be due to the fact that there is insufficient information on those regions such as promoter and transcription start site or the motif is too small to be detected with confidence. It is known that variation on the utilization of mRNA template contributes to the localization of protein. One example is the invertase gene in *S. cerevisiae* which uses the same locus to encode cytoplasmic and extracellular invertase depending on the translational start site. Not all proteins are secreted by the conventional secretory pathways. Those proteins are processed in the cell and exported out of the cell by ATP-Binding Cassette transporters, suggesting ABC transporters play an important role in the export of some secreted proteins. A large scale EST data analysis of the plant pathogen, *Phytophthora infestans*, using exploitation of bioinformatic approaches and functional assays was successful to identifying novel pathogen effector proteins (Torto et al., 2003). It is clear that genomic data, whether annotated or not, has to be verified using experimental approaches. The first version of a signal sequence trap (SST) method utilized a vector construct that had a cloning site, an epitope, and a transmembrane domain protein as a translational fusion (Tashiro et al., 1993). If DNA inserted in the cloning site provided a start codon and signal peptide, the protein was expressed. The presence of signal peptide brings the translation product to the plasma membrane where the protein is trapped by the transmembrane domain with the epitope exposed outside of the cell. The original method possessed some drawbacks. First, the address of the positive clone needed to be confirmed by performing individual screens since a pooling strategy was used for initial screening. Second, if a very large and complex library was to be screened, it required a deep coverage transfection, a labor intensive pooling, and readdressing procedure. Reporter gene strategies are attractive because phenotypes can

be easily identified with visual inspection such as chromogenic reaction using phosphatase reaction (Chen and Leder, 1999). Alkaline phosphatase has been used extensively as a reporter because it is robust and a wide variety of substrates are available. However, reporter-based strategies also suffer when a complex and large library needs to be screened. Thus, experimental systems that combine the ease of screening and capability of handling a complex library are needed. Furthermore, selection systems based on nutritional requirement or drug-sensitivity are preferred. In this respect, the yeast invertase-based system is attractive (Klein et al., 1996). The yeast invertase, SUC2, encodes both cytoplasmic and extracellular forms, depending on the translational start site used. The invertase-based secretory protein selection system uses the cytoplasmic form of invertase gene without a start codon. Foreign DNA is inserted at the 5'-end of invertase. If the foreign DNA provides a start codon and signal peptide, the fusion protein will be exported and sucrose is utilized as a carbon source. The recipient mutant strain is deficient in *suc2*, and can not grow on medium containing sucrose as a sole carbon source. Besides the use of invertase as a selectable marker, a chromogenic invertase assay is easy to perform and provides quantitative data. The assay can be exploited to confirm the presence of secreted invertase in growth medium. Thus, invertase as both reporter gene and selectable marker provides an efficient method for identifying secretory proteins using yeast as a proxy host.

E. coli also has been used as a host to screen and identify secreted proteins from eukaryotic cells, and functions based on the observation that signal peptides in both eukaryote and gram negative bacteria are often similar (Tan et al., 2003). Translational fusion proteins are excreted into the periplasmic space where the translational fusion

protein confers antibiotic resistance and chromogenic enzymatic activity. Since prokaryotes have no cellular organelles and there is some evidence on the inter-specific statistical variation on signal peptide functionality, such a system may not be most appropriate for screening eukaryotic sources.

Microbial cell surface display systems are quite attractive and efficient to investigate in vitro protein-protein or protein-ligand interactions. Deep coverage libraries can be made with ease, and the screening process is often efficient. The yeast cell surface display was used to generate scFv antibody library and affinity maturation (Boder and Wittrup, 1997; Feldhaus et al., 2003). scFv antibody library was made and screened against antigens. The yeast display system allows separation of positive cells by techniques such as magnetic beads and a high capacity fluorescence activate cell sorter (FACS). Either of these methods or the combination of both allows screening of very deep coverage libraries by reducing the number of cells to be screened at each step. Since the system uses eukaryotic cells, the proteins that are secreted and displayed are under strict quality control during the translocation from endoplasmic reticulum and Golgi apparatus, there is less concern about protein activity or confirmation of displayed proteins even though there may be differences in protein adornment and processing. Thus, the combination of efficient selection and enrichment strategies such as FACS will enable genome-wide screening of target molecules, and facilitate functional characterization of genes with unknown function. Also, the development of an efficient genome-wide screening tool will contribute to genome annotation, discovery of new genes, and functional motifs.

Application of microarray for molecular localization and functional genomics

DNA microarrays have revolutionized molecular biology and genomics sciences (Schemm et al., 1995). The ability to measure expression levels of thousands of genes at a time has made it possible to identify genes that are involved in certain conditions such as cancer classification, developmental processes, and responses to external perturbations. Typically, DNA microarrays are used for hybridization of targets and probes, probing quantitative expression level of genes or qualitative presence or absence of targets in a complex mixture of nucleic acids. However, the microarray format is also suited to probe other interactions such as DNA-protein interactions. Microarrays can be constructed using proteins, and used to detect interaction with other proteins, nucleic acids or other molecules (Ptacek et al., 2005; Zhu et al., 2000).

It is well-known that measurement of steady-state transcript levels is not the sole indicator of the activity of the gene and gene product. There are many factors that affect the activity of genes and their protein products. The localization of transcript message and protein products provides an important clue as to function. This is well documented during the developmental stages of a variety of organisms. Even in unicellular organisms such as *S. cerevisiae*, mRNA and protein localization plays an important role such as ASH1 gene that functions in the budding process (Shepard et al., 2003).

While proteins that are destined to be cytosolic are translated in the free cytosolic ribosomes, proteins that are to be transported out of the cell are translated on ribosomes that are bound to ER membranes. However, there are exceptions. There are two mechanisms by which protein translocation into ER lumen occurs other than exportation by ABC transporters. One is a co-translational process, and the other is a post-

translational process. The former mechanism is characterized by the recognition of the signal peptide by signal recognition particle. The nascent protein/mRNA/ribosome complex is brought to the ER membrane by signal recognition particle (SRP) with arrested or slowed translation. As translation proceeds, the protein is inserted into the ER lumen and processed. Post-translational translocation is mediated by Sec complex including Sec61 protein and BiP with at least two distinct steps: substrate binding to the channel and translocation (Rapoport et al., 1999). It is postulated that mode of translocation is determined by the strength of signal peptide such as hydrophobicity (Ng et al., 1996). Cell biology techniques can separate cellular organelles, and is also possible to fractionate ER membrane and cytosolic fraction along with ribosomes and mRNA into different fractions. Mitochondria targeted genes have been investigated using similar approaches (Marc et al., 2002). Cellular contents are separated using density gradient ultra-centrifugation or differential sediment centrifugation. In combination with high throughput microarray, these techniques were used to discriminate transcripts encoding membrane-bound and secreted proteins from transcripts encoding cytosolic proteins in budding yeast and human (Diehn et al., 2000; Toyoda et al., 2003). Many transcripts were found in both fractions, which were speculated not to be erroneous but due to unknown biological reasons. Using mammalian cells, a similar approach confirmed an existing catalog of secreted and membrane-bound protein, corrected prediction, and/or found novel genes (Toyoda et al., 2003). Thus, exploitation of both existing and new technologies will likely add significantly to our knowledge of the rice blast transcriptome.

Conclusion

The genome sequencing of *M. grisea* paved the avenue for functional analysis of its genomic content. Bioinformatics is an essential part of genomic studies, providing a roadmap for experimental investigations. Recent reports on large scale EST analysis suggest that the complexity of transcriptome might be more than we expected. Discovery of micro RNAs and non-coding RNAs in other organisms strongly suggest that similar findings may be forthcoming in filamentous fungi, too. Also, the detection and characterization of these RNA might lead to the discovery of novel regulatory roles. Thus, functional genomics based on expression analysis will play an essential part of future research. Budding yeast has served as a workhorse of functional genomics in a variety of capacities, leading to the development of technology and transfer of knowledge to other systems. Filamentous fungi are much more diverse, and exhibit a broad range of phenotypes that are not present in the budding yeast such as infection-related development, diverse sexual structures, and diverse ecological niches they occupy. The main bottle neck remains the development of strategies for high-throughput functional analysis of gene content. Progress of functional genomics in the rice blast fungus has been prolific, owing to the generation of genome sequencing, availability of high-density microarrays, development of large scale transformation system such as ATMT, and a large pool of insertional mutagenesis strains. Future research will include expression-based functional analysis. Microarray remains a major platform for functionally linking to sequence data. Expression profiling will provide a valuable checkpoint for prioritizing functional analysis through over-expression, targeted knockout, and possibly through RNAi-based knockdown approaches. However, there are other opportunities to be

explored such as the generation of non-homologous end joining (NHEJ)-deficient strains for high efficiency homologous recombination for targeted mutagenesis as in *N. crassa*, protein over-expression using cDNA library resources for functional studies, and large scale protein-protein interaction studies for identification of host-pathogen interactions and elucidation of pathogenicity mechanisms inside the pathogen.

References

1. Adams,D.G., Sachs,N.A., and Vaillancourt,R.R. (2002). Phosphorylation of the stress-activated protein kinase, MEKK3, at serine 166. *Arch Biochem Biophys* 407, 103-116.
2. Adams,M.D., Kelley,J.M., Gocayne,J.D., Dubnick,M., Polymeropoulos,M.H., Xiao,H., Merril,C.R., Wu,A., Olde,B., Moreno,R.F., and . (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252, 1651-1656.
3. Armstrong,M.R., Whisson,S.C., Pritchard,L., Bos,J.I., Venter,E., Avrova,A.O., Rehmany,A.P., Bohme,U., Brooks,K., Cherevach,I., Hamlin,N., White,B., Fraser,A., Lord,A., Quail,M.A., Churcher,C., Hall,N., Berriman,M., Huang,S., Kamoun,S., Beynon,J.L., and Birch,P.R. (2005). An ancestral oomycete locus contains late blight avirulence gene *Avr3a*, encoding a protein that is recognized in the host cytoplasm. *Proc. Natl Acad. Sci U. S. A* 102, 7766-7771.
4. Austin,M.J., Muskett,P., Kahn,K., Feys,B.J., Jones,J.D.G., and Parker,J.E. (2002). Regulatory role of SGT1 in early R gene-mediated plant defenses. *Science* 295, 2077-2080.
5. Beck,T. and Hall,M.N. (1999). The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors. *Nature* 402, 689-692.

6. Bendtsen,J.D., Nielsen,H., von Heijne,G., and Brunak,S. (2004). Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* *340*, 783-795.
7. Birch,P.R.J., Rehmany,A.P., Pritchard,L., Kamoun,S., and Beynon,J.L. (2006). Trafficking arms: oomycete effectors enter host plant cells. *Trends Microbiol* *14*, 8-11.
8. Boder,E.T. and Wittrup,K.D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotech* *15*, 553-557.
9. Bonaldo,M.F., Lennon,G., and Soares,M.B. (1996). Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Res.* *6*, 791-806.
10. Bryan,G.T., Wu,K.S., Farrall,L., Jia,Y., Hershey,H.P., McAdams,S.A., Faulk,K.N., Donaldson,G.K., Tarchini,R., and Valent,B. (2000). tA single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene Pi-ta. *Plant Cell* *12*, 2033-2046.
11. Capecchi,M.R. (1989). Alternating the genome by homologous recombination. *Science* *244*, 1288-1292.
12. Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y. (2000). Normalization and

- subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* *10*, 1617-1630.
13. Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y. (1998). Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. *Proc. Natl. Acad. Sci U. S. A* *95*, 520-524.
 14. Catanzariti,A.M., Dodds,P.N., Lawrence,G.J., Ayliffe,M.A., and Ellis,J.G. (2006). Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* *18*, 243-256.
 15. Chen,H. and Leder,P. (1999). A new signal sequence trap using alkaline phosphatase as a reporter. *Nucl. Acids Res.* *27*, 1219-1222.
 16. Cullen,P.J. and Sprague,G.F., Jr. (2000). Glucose depletion causes haploid invasive growth in yeast. *Proc. Natl. Acad. Sci U. S. A* *97*, 13619-13624.
 17. Dangl,J.L. and Jones,J.D.G. (2001). Plant pathogens and integrated defence responses to infection. *Nature* *411*, 826-833.
 18. Das,M., Harvey,I., Chu,L.L., Sinha,M., and Pelletier,J. (2001). Full-length cDNAs: more than just reaching the ends. *Physiol Genomics* *6*, 57-80.

19. de Groot,M.J., Bundock,P., Hooykaas,P.J., and Beijersbergen,A.G. (1998). *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. Nat Biotechnol. 16, 839-842.
20. Dean,R.A. (1997). Signal pathways and appressorium morphogenesis. Annu. Rev. Phytopathol. 35, 211-234.
21. Dean,R.A., Talbot,N.J., Ebbole,D.J., Farman,M.L., Mitchell,T.K., Orbach,M.J., Thon,M., Kulkarni,R., Xu,J.R., Pan,H., Read,N.D., Lee,Y.H., Carbone,I., Brown,D., Oh,Y.Y., Donofrio,N., Jeong,J.S., Soanes,D.M., Djonovic,S., Kolomiets,E., Rehmeyer,C., Li,W., Harding,M., Kim,S., Lebrun,M.H., Bohnert,H., Coughlan,S., Butler,J., Calvo,S., Ma,L.J., Nicol,R., Purcell,S., Nusbaum,C., Galagan,J.E., and Birren,B.W. (2005). The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature 434, 980-986.
22. DeZwaan,T.M., Carroll,A.M., Valent,B., and Sweigard,J.A. (1999). *Magnaporthe grisea* Pth11p is a novel plasma membrane protein that mediates appressorium differentiation in response to inductive surface cues. Plant Cell 11, 2012-2030.
23. Diehn,M., Eisen,M.B., Botstein,D., and Brown,P.O. (2000). Large-scale identification of secreted and membrane-associated gene products using DNA microarrays. Nat Genet 25, 58-62.

24. Dumaz,N. and Marais,R. (2005). Integrating signals between cAMP and the RAS/RAF/MEK/ERK signalling pathways. *FEBS J* 272, 3491-3504.
25. Eby,J.J., Holly,S.P., van Drogen,F., Grishin,A.V., Peter,M., Drubin,D.G., and Blumer,K.J. (1998). Actin cytoskeleton organization regulated by the PAK family of protein kinases. *Curr. Biol* 8 , 967-970.
26. Edery,I., Chu,L.L., Sonenberg,N., and Pelletier,J. (1995). An efficient strategy to isolate full-length cDNAs based on an mRNA cap retention procedure (CAPture). *Mol. Cell. Biol.* 15, 3363-3371.
27. Emmett,R.W. and Parbery,D.G. (1975). Appressoria. *Annu. Rev. Phytopathol.* 13, 147-165.
28. Fairhead,C., Llorente,B., Denis,F., Soler,M., and Dujon,B. (1996). New vectors for combinatorial deletions in yeast chromosomes and for gap-repair cloning using 'split-marker' recombination. *Yeast* 12, 1439-1457.
29. Fanger,G.R., Widmann,C., Porter,A.C., Sather,S., Johnson,G.L., and Vaillancourt,R.R. (1998). 14-3-3 proteins interact with specific MEK kinases. *J Biol Chem.* 273, 3476-3483.
30. Feldhaus,M.J., Siegel,R.W., Opresko,L.K., Coleman,J.R., Feldhaus,J.M.W., Yeung,Y.A., Cochran,J.R., Heinzelman,P., Colby,D., Swers,J., Graff,C.,

- Wiley,H.S., and Wittrup,K.D. (2003). Flow-cytometric isolation of human antibodies from a nonimmune *Saccharomyces cerevisiae* surface display library. *Nat Biotech* 21, 163-170.
31. Flaishman,M.A. and Kolattukudy,P.E. (1994). Timing of fungal invasion using host's ripening hormone as a signal. *Proc. Natl. Acad. Sci U. S. A* 91, 6579-6583.
 32. Fudal,I., Bohnert,H.U., Tharreau,D., and Lebrun,M.H. (2005). Transposition of MINE, a composite retrotransposon, in the avirulence gene ACE1 of the rice blast fungus *Magnaporthe grisea*. *Fungal Genet. Biol.*
 33. Galagan,J.E. (2003). The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422, 859-868.
 34. Gelperin,D., Weigle,J., Nelson,K., Roseboom,P., Irie,K., Matsumoto,K., and Lemmon,S. (1995). 14-3-3 proteins: potential roles in vesicular transport and Ras signaling in *Saccharomyces cerevisiae*. *PNAS* 92, 11539-11543.
 35. Gilbert,R.D., Johnson,A.M., and Dean,R.A. (1996). Chemical signals responsible for appressorium formation in the rice blast fungus *Magnaporthe grisea*. *Physiol Mol Plant P* 48, 335-346.

36. Gimeno,C.J., Ljungdahl,P.O., Styles,C.A., and Fink,G.R. (1992). Unipolar cell divisions in the yeast *S. cerevisiae* lead to filamentous growth: regulation by starvation and RAS. *Cell* 68, 1077-1090.
37. Goff,S. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296, 92-100.
38. Goff,S.A. (1999). Rice as a model for cereal genomics. *Curr. Opin. Plant Biol* 2, 86-89.
39. Graves,P.R., Lovly,C.M., Uy,G.L., and Piwnica-Worms,H. (2001). Localization of human Cdc25C is regulated both by nuclear export and 14-3-3 protein binding. *Oncogene* 20, 1839-1851.
40. Hamer,L., Adachi,K., Montenegro-Chamorro,M.V., Tanzer,M.M., Mahanty,S.K., Lo,C., Tarpey,R.W., Skalchunes,A.R., Heiniger,R.W., Frank,S.A., Darveaux,B.A., Lampe,D.J., Slater,T.M., Ramamurthy,L., DeZwaan,T.M., Nelson,G.H., Shuster,J.R., Woessner,J., and Hamer,J.E. (2001). Gene discovery and gene function assignment in filamentous fungi. *Proc. Natl Acad. Sci U. S. A* 98, 5110-5115.
41. Hiller,N.L., Bhattacharjee,S., van Ooij,C., Liolios,K., Harrison,T., Lopez-Estrano,C., and Haldar,K. (2004). A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* 306, 1934-1937.

42. Hoch,H.C., Staples,R.C., Whitehead,B., Commeau,J., and Wolf,E.D. (1987). Signalling for growth orientation and cell differentiation by surface topology in *Uromyces*. *Science* 235, 1659-1662.
43. Holly,S.P. and Blumer,K.J. (1999). PAK-family kinases regulate cell and actin polarization throughout the cell cycle of *Saccharomyces cerevisiae*. *J. Cell Biol.* 147, 845-856.
44. Hurtado,C.A.R. and Rachubinski,R.A. (2002). YIBMH1 encodes a 14-3-3 protein that promotes filamentous growth in the dimorphic yeast *Yarrowia lipolytica*. *Microbiology* 148, 3725-3735.
45. Igarashi,D., Ishida,S., Fukazawa,J., and Takahashi,Y. (2001). 14-3-3 proteins regulate intracellular localization of the bZIP transcriptional activator RSG. *Plant Cell* 13, 2483-2497.
46. Irie,K., Gotoh,Y., Yashar,B.M., Errede,B., Nishida,E., and Matsumoto,K. (1994). Stimulatory effects of yeast and mammalian 14-3-3 proteins on the Raf protein kinase. *Science* 265, 1716-1719.
47. Jenczmionka,N.J. and Schafer,W. (2005). The Gpmk1 MAP kinase of *Fusarium graminearum* regulates the induction of specific secreted enzymes. *Curr. Genet* 47, 29-36.

48. Jia, Y., McAdams, S.A., Bryan, G.T., Hershey, H.P., and Valent, B. (2000). Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *EMBO J.* *19*, 4004-4014.
49. Kahmann, R. and Basse, C. (2001). Fungal gene expression during pathogenesis-related development and host plant colonization. *Curr. Opin. Microbiol* *4*, 374-380.
50. Kang, S., Lebrun, M.H., Farrall, L., and Valent, B. (2001). Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. *Mol. Plant Microbe Interact.* *14*, 671-674.
51. Keen, N.T. (1990). Gene-for-gene complementarity in plant-pathogen interactions. *Annu. Rev Genet* *24*, 447-463.
52. Khang, C.H., Park, S.Y., Lee, Y.H., and Kang, S. (2005). A dual selection based, targeted gene replacement tool for *Magnaporthe grisea* and *Fusarium oxysporum*. *Fungal Genet Biol* *42*, 483-492.
53. Khush, G.S. (2005). What it will take to feed 5.0 billion rice consumers in 2030. *Plant Mol Biol* *59*, 1-6.

54. Klein,R.D., Gu,Q., Goddard,A., and Rosenthal,A. (1996). Selection for genes encoding secreted proteins and receptors. *Proc. Natl. Acad. Sci U. S. A* *93*, 7108-7113.
55. Kraus,P.R., Hofmann,A.F., and Harris,S.D. (2002). Characterization of the *Aspergillus nidulans* 14-3-3 homologue, ArtA. *FEMS Microbiol Lett.* *210*, 61-66.
56. Kulkarni,R.D. and Dean,R.A. (2004). Identification of proteins that interact with two regulators of appressorium development, adenylyate cyclase and cAMP-dependent protein kinase A, in the rice blast fungus *Magnaporthe grisea*. *Mol Genet. Genomics* *270*, 497-508.
57. Kulkarni,R.D., Thon,M.R., Pan,H., and Dean,R.A. (2005). Novel G-protein-coupled receptor-like proteins in the plant pathogenic fungus *Magnaporthe grisea*. *Genome Biol* *6*, R24.
58. Lee,Y.H. and Dean,R.A. (1993). cAMP regulates infection structure formation in the plant pathogenic fungus *Magnaporthe grisea*. *Plant Cell* *5*, 693-700.
59. Leung,H., Lehtinen,U., Karjalainen,R., Skinner,D., Tooley,P., Leong,S., and Ellingboe,A. (1990). Transformation of the rice blast fungus *Magnaporthe grisea* to hygromycin B resistance. *Curr. Genet.* *17*, 409-411.

60. Li,L., Xue,C., Bruno,K., Nishimura,M., and Xu,J.R. (2004). Two PAK kinase genes, CHM1 and MST20, have distinct functions in *Magnaporthe grisea*. Mol Plant Microbe Interact. 17, 547-556.
61. Lopez-Girona,A., Furnari,B., Mondesert,O., and Russell,P. (1999). Nuclear localization of Cdc25 is regulated by DNA damage and a 14-3-3 protein. Nature 397, 172-175.
62. Mansour,S.L., Thomas,K.R., and Capecchi,M.R. (1988). Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature 336, 348-352.
63. Marc,P., Margeot,A., Devaux,F., Blugeon,C., Corral-Debrinski,M., and Jacq,C. (2002). Genome-wide analysis of mRNAs targeted to yeast mitochondria. EMBO Rep. 3, 159-164.
64. Maruyama,K. and Sugano,S. (1994). Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. Gene 138, 171-174.
65. Mitchell,T.K. and Dean,R.A. (1995). The cAMP-dependent protein kinase catalytic subunit is required for appressorium formation and pathogenesis by the rice blast pathogen *Magnaporthe grisea*. Plant Cell 7, 1869-1878.

66. Ng,D.T., Brown,J.D., and Walter,P. (1996). Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J Cell Biol* *134*, 269-278.
67. Nicholson,L.R. (1996). Adhesion of fungal propagules. In *Histology, ultrastructure and molecular cytology of plant-micorbe interactions*, M.Nicole and V.Gianinazzi-Pearson, eds. Kluwer academic publusher), pp. 117-134.
68. Ninomiya,Y., Suzuki,K., Ishii,C., and Inoue,H. (2004). Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. *Proc. Natl Acad. Sci U. S. A* *101*, 12248-12253.
69. Ohara,O. and Temple,G. (2001). Directional cDNA library construction assisted by the in vitro recombination reaction. *Nucleic Acids Res.* *29*, E22.
70. Orbach,M.J., Farrall,L., Sweigard,J.A., Chumley,F.G., and Valent,B. (2000). A telomeric avirulence gene determines efficacy for the rice blast resistance gene *Pi-ta*. *Plant Cell* *12*, 2019-2032.
71. Ospina-Giraldo,M.D., Mullins,E., and Kang,S. (2003). Loss of function of the *Fusarium oxysporum* SNF1 gene reduces virulence on cabbage and *Arabidopsis*. *Curr. Genet* *44*, 49-57.

72. Pan,X. and Heitman,J. (1999). Cyclic AMP-dependent protein kinase regulates pseudohyphal differentiation in *Saccharomyces cerevisiae*. *Mol Cell Biol* *19*, 4874-4887.
73. Park,G., Xue,C., Zheng,L., Lam,S., and Xu,J.R. (2002). MST12 regulates infectious growth but not appressorium formation in the rice blast fungus *Magnaporthe grisea*. *Mol Plant Microbe Interact.* *15*, 183-192.
74. Ptacek,J., Devgan,G., Michaud,G., Zhu,H., Zhu,X., Fasolo,J., Guo,H., Jona,G., Breikreutz,A., Sopko,R., McCartney,R.R., Schmidt,M.C., Rachidi,N., Lee,S.J., Mah,A.S., Meng,L., Stark,M.J., Stern,D.F., De Virgilio,C., Tyers,M., Andrews,B., Gerstein,M., Schweitzer,B., Predki,P.F., and Snyder,M. (2005). Global analysis of protein phosphorylation in yeast. *Nature* *438*, 679-684.
75. Rapoport,T.A., Matlack,K.E., Plath,K., Misselwitz,B., and Staack,O. (1999). Posttranslational protein translocation across the membrane of the endoplasmic reticulum. *Biol Chem.* *380*, 1143-1150.
76. Reid,R.J., Lisby,M., and Rothstein,R. (2002). Cloning-free genome alterations in *Saccharomyces cerevisiae* using adaptamer-mediated PCR. *Methods Enzymol.* *350*, 258-277.

77. Rho,H.S., Kang,S., and Lee,Y.H. (2001). *Agrobacterium tumefaciens*-mediated transformation of the plant pathogenic fungus, *Magnaporthe grisea*. *Mol Cells* 12, 407-411.
78. Roberts,R.L. and Mosch,H.U. (1997). 14-3-3 proteins are essential for RAS/MAPK cascade signaling during pseudohyphal development in *S. cerevisiae*. *Cell* 89, 1055-1065.
79. Roth,D., Birkenfeld,J., and Betz,H. (1999). Dominant-negative alleles of 14-3-3 proteins cause defects in actin organization and vesicle targeting in the yeast *Saccharomyces cerevisiae*. *FEBS Lett.* 460, 411-416.
80. Rudd,S. (2003). Expressed sequence tags: alternative or complement to whole genome sequences? *Trends Plant Sci* 8, 321-329.
81. Schafer,W. (1994). Molecular mechanisms of fungal pathogenicity to plants. *Annu. Rev Phytopathol.* 32, 461-477.
82. Schena,M., Shalon,D., Davis,R.W., and Brown,P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467-470.

83. Schmidt, W.M. and Mueller, M.W. (1999). CapSelect: a highly sensitive method for 5' CAP-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. *Nucl. Acids Res.* 27, e31.
84. Sesma, A. and Osbourn, A.E. (2004). The rice blast pathogen undergoes developmental processes typical of root-infecting fungi. *Nature* 431, 582-586.
85. Shao, F., Golstein, C., Ade, J., Stoutemyer, M., Dixon, J.E., and Innes, R.W. (2003). Cleavage of *Arabidopsis* PBS1 by a bacterial type III effector. *Science* 301, 1230-1233.
86. Shepard, K.A., Gerber, A.P., Jambhekar, A., Takizawa, P.A., Brown, P.O., Herschlag, D., DeRisi, J.L., and Vale, R.D. (2003). Widespread cytoplasmic mRNA transport in yeast: Identification of 22 bud-localized transcripts using DNA microarray analysis. *Proc. Natl. Acad. Sci U. S. A* 100, 11429-11434.
87. Skinner, W., Keon, J., and Hargreaves, J. (2001). Gene information for fungal plant pathogens from expressed sequences. *Curr. Opin. Microbiol* 4, 381-386.
88. Sweigard, J.A., Chumley, F.G., and Valent, B. (1992). Disruption of a *Magnaporthe grisea* cutinase gene. *Molecular Genetics and Genomics (Historical Archive)* 232, 183-190.

89. Sweigard, J.A. and Ebbole, D.J. (2001). Functional analysis of pathogenicity genes in a genomics world. *Curr. Opin. Microbiol* 4, 387-392.
90. Tan, R., Jiang, X., Jackson, A., Jin, P., Yang, J., Lee, E., Duggan, B., Stuve, L.L., and Fu, G.K. (2003). *E. coli* selection of human genes encoding secreted and membrane proteins based on cDNA fusions to a leaderless β -lactamase reporter. *Genome Res.* 13, 1938-1943.
91. Tashiro, K., Tada, H., Heilker, R., Shirozu, M., Nakano, T., and Honjo, T. (1993). Signal sequence trap: a cloning strategy for secreted proteins and type I membrane proteins. *Science* 261, 600-603.
92. The FANTOM Consortium, Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impiombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Gatta, G.D., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminiecki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasaki, Y., Kelso, J., Kitamura, H.,

Kitano,H., Kollias,G., Krishnan,S.P.T., Kruger,A., Kummerfeld,S.K.,
Kurochkin,I.V., Lareau,L.F., Lazarevic,D., Lipovich,L., Liu,J., Liuni,S.,
McWilliam,S., Babu,M.M., Madera,M., Marchionni,L., Matsuda,H.,
Matsuzawa,S., Miki,H., Mignone,F., Miyake,S., Morris,K., Mottagui-Tabar,S.,
Mulder,N., Nakano,N., Nakauchi,H., Ng,P., Nilsson,R., Nishiguchi,S.,
Nishikawa,S., Nori,F., Ohara,O., Okazaki,Y., Orlando,V., Pang,K.C., Pavan,W.J.,
Pavesi,G., Pesole,G., Petrovsky,N., Piazza,S., Reed,J., Reid,J.F., Ring,B.Z.,
Ringwald,M., Rost,B., Ruan,Y., Salzberg,S.L., Sandelin,A., Schneider,C.,
Schonbach,C., Sekiguchi,K., Semple,C.A.M., Seno,S., Sessa,L., Sheng,Y.,
Shibata,Y., Shimada,H., Shimada,K., Silva,D., Sinclair,B., Sperling,S., Stupka,E.,
Sugiura,K., Sultana,R., Takenaka,Y., Taki,K., Tammoja,K., Tan,S.L., Tang,S.,
Taylor,M.S., Tegner,J., Teichmann,S.A., Ueda,H.R., van Nimwegen,E.,
Verardo,R., Wei,C.L., Yagi,K., Yamanishi,H., Zabarovsky,E., Zhu,S., Zimmer,A.,
Hide,W., Bult,C., Grimmond,S.M., Teasdale,R.D., Liu,E.T., Brusic,V.,
Quackenbush,J., Wahlestedt,C., Mattick,J.S., Hume,D.A., RIKEN Genome
Exploration Research Group and Genome Science Group (Genome Network
Project Core Group), Kai,C., Sasaki,D., Tomaru,Y., Fukuda,S., Kanamori-
Katayama,M., Suzuki,M., Aoki,J., Arakawa,T., Iida,J., Imamura,K., Itoh,M.,
Kato,T., Kawaji,H., Kawagashira,N., Kawashima,T., Kojima,M., Kondo,S.,
Konno,H., Nakano,K., Ninomiya,N., Nishio,T., Okada,M., Plessy,C., Shibata,K.,
Shiraki,T., Suzuki,S., Tagami,M., Waki,K., Watahiki,A., Okamura-Oho,Y.,
Suzuki,H., Kawai,J., and Hayashizaki,Y. (2005). The Transcriptional landscape of
the mammalian genome. *Science* 309, 1559-1563.

93. Thines,E., Weber,R.W.S., and Talbot,N.J. (2000). MAP kinase and protein kinase A-dependent mobilization of triacylglycerol and glycogen during appressorium turgor generation by *Magnaporthe grisea*. *Plant Cell* *12*, 1703-1718.
94. Tonukari,N.J., Scott-Craig,J.S., and Walton,J.D. (2000). The *Cochliobolus carbonum* SNF1 gene is required for cell wall-degrading enzyme expression and virulence on maize. *Plant Cell* *12*, 237-248.
95. Torto,T.A., Li,S., Styer,A., Huitema,E., Testa,A., Gow,N.A., van West,P., and Kamoun,S. (2003). EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*. *Genome Res.* *13*, 1675-1685.
96. Toyoda,N., Nagai,S., Terashima,Y., Motomura,K., Haino,M., Hashimoto,S.i., Takizawa,H., and Matsushima,K. (2003). Analysis of mRNA with microsomal fractionation using a SAGE-Based DNA microarray system facilitates identification of the genes encoding secretory proteins. *Genome Res.* *13*, 1728-1736.
97. Valent,B. and Chumley,F.G. (1991). Molecular genetic analysis of the rice blast fungus *Magnaporthe grisea*. *Annu. Rev. Phytopathol.* *29*, 443-467.

98. Vasara,T., Keranen,S., Penttila,M., and Saloheimo,M. (2002). Characterisation of two 14-3-3 genes from *Trichoderma reesei*: interactions with yeast secretory pathway components. *Biochim Biophys Acta*. *1590*, 27-40.
99. Wach,A., Brachat,A., Pohlmann,R., and Philippsen,P. (1994). New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast* *10*, 1793-1808.
100. Walton,J.D. (1994). Deconstructing the cell wall. *Plant Physiol* *104*, 1113-1118.
101. Wu,S.C., Ham,K.S., Darvill,A.G., and Albersheim,P. (1997). Deletion of two endo-beta-1,4-xylanase genes reveals additional isozymes secreted by the rice blast fungus. *Mol Plant Microbe Interact*. *10*, 700-708.
102. Xu,J.R., Urban,M., Sweigard,J.A., and Hamer,J.E. (1997). The CPKA gene of *Magnaporthe grisea* is essential for appressorial penetration. *Mol Plant Microbe Interact*. *10*, 187-194.
103. Yu,J. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* *296*, 79-92.
104. Zhao,X., Kim,Y., Park,G., and Xu,J.R. (2005). A mitogen-activated protein kinase cascade regulating infection-related morphogenesis in *Magnaporthe grisea*. *Plant Cell* *17*, 1317-1329.

105. Zhu,H., Klemic,J.F., Chang,S., Bertone,P., Casamayor,A., Klemic,K.G., Smith,D., Gerstein,M., Reed,M.A., and Snyder,M. (2000). Analysis of yeast protein kinases using protein chips. *Nat Genet* 26, 283-289.

106. Zhu,Y.Y., Machleder,E.M., Chenchik,A., Li,R., and Siebert,P.D. (2001). Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *Biotechniques* 30 , 892-897.

Chapter Two

A novel strategy for cDNA library generation using SMART (Switching Mechanism At 5' end of RNA Template) and recombinational cloning systems

Jun Seop Jeong and Ralph A. Dean

Center for Integrated Fungal Research (CIFR), North Carolina State University
Raleigh, NC 27695

Abstract

cDNA libraries represent a central part of genomic studies. They serve as a reagent for large scale EST analysis, gene model prediction, expression profiling, and protein expression. Large scale cDNA sequencing projects have facilitated discovery of novel genes. Also, cDNA libraries generated from condition-specific, tissue-specific or certain developmental processes have provided insight into physiology of the biological process under investigation. cDNA library construction and EST sequencing can be a cost-effective way for genome investigations. However, cDNA libraries typically suffer redundancy that is inherent in RNA populations from which libraries are made. Thus, cDNA normalization is frequently employed. To aid annotation of genome sequences, recent studies have been focused on the construction of full length cDNA libraries. However, construction of a full length cDNA library often requires demanding techniques and a large amount of starting materials. SMART (Switching Mechanism At 5' end of RNA Template) technology offers a straightforward approach for creating full length enriched cDNA libraries. In this work, we combined a recombination cloning strategy with SMART technology to construct a cDNA library from the rice blast fungus, *Magnaporthe grisea*. We also demonstrated that the cDNA library could be readily normalized. Sequence analysis suggested that the library was enriched for full length cDNAs, and contained cDNA clones for novel transcripts. Thus, our method provides an efficient way of generating full length enriched normalized cDNA libraries.

Introduction

Full length cDNA is invaluable for numerous molecular biology and genomic studies (Das et al., 2001). Early efforts to find genes in complex genomes were greatly facilitated by large scale EST (expressed sequence tag) sequencing of cDNA libraries (Adams et al., 1991). Subsequently, EST sequencing has greatly aided identification of genes involved in certain stages of development for large scale whole genome sequencing projects (Bonaldo et al., 1996). It is a cost-effective way of identifying genes in complex genomes. However, large scale EST sequence typically suffered from problems with redundancy that is inherent in disproportional mRNA population in cells, and results in biased sequencing of abundant messages. To this end, cDNA libraries are often normalized to increase the complexity of the library and to facilitate the finding of new genes (Bonaldo et al., 1996; Soares et al., 1994) .

In addition to the importance of cDNA library generation for EST sequencing, cDNA libraries are indispensable reagents for functional genomics. Since comparative analysis of EST data may reflect the physiological condition under investigation, occurrence of genes in certain conditions may provide clues to the function, leading to prioritization for functional analysis. Often, cDNA is the template for protein expression studies. cDNA clones with known sequence information can be directly used for protein expression. Recently, large scale proteomic approaches have been undertaken using cDNA libraries (Lamesch et al., 2004; Luan et al., 2004; Rual et al., 2004). Thus, cDNA generation and EST sequencing data are the bridge between structural and functional genomics.

More recent efforts have been focused on constructing full length cDNA libraries. Full length cDNA offers several advantages. First, full length cDNA can contribute to the generation of more accurate gene models or defining transcription units. Large scale EST analysis from full length cDNA libraries has indicated that transcriptional activity of a genome can be at least an order of magnitude larger than *in silico* gene prediction (The FANTOM Consortium et al., 2005). The data also indicate that variation arising from alternative splicing, use of alternative uses of promoter and polyadenylation sites is greatly underestimated. Furthermore, full length cDNA is more useful for protein expression studies. Recent ORFeome approaches have shown the utility of cDNA libraries for large scale protein expression and protein-protein interaction studies (Reboul et al., 2003; Rual et al., 2005). Thus, deep coverage full length cDNA libraries are essential for functional genomic and proteomic studies. However, a difficulty with the generation of full length cDNA libraries is the lack of defined 5' sequence compared to poly-A tailed 3' end. cDNA library construction exploits the 3' poly-A tail. Often, full length cDNA generation is limited by premature termination of reverse transcriptase (RT) (Carninci et al., 1998). cDNA library generation has been traditionally done with first strand synthesis RT and second strand synthesis using bacterial or phage origin DNA polymerases such as *E. coli* DNA polymerase I and T4 DNA polymerase (Gubler and Hoffman, 1983). Subsequently, several steps of ligation and purification are usually followed, producing limiting amounts of cDNA for library generation. Recently, a strategy for cDNA library generation was developed exploiting an intrinsic enzymatic property of RT (Schmidt and Mueller, 1999; Zhu et al., 2001). MMLV RT has the enzymatic property of terminal transferase. SMART cDNA technology obviates the need

for adaptor ligation. Terminal transferase activity allows the addition of a switching oligo, so that the switching oligo can be used to synthesize the second strand. Switching is more efficient with full length mRNA which has a specialized cap structure (Schmidt and Mueller, 1999). Also, cDNA with premature termination of RT would be amplified in a linear fashion compared to the exponential amplification of full length cDNA with an intact 3' and 5' end. With the known 5' and 3'-ends, it is possible to amplify the cDNA by PCR, again obviating the chance of limiting amounts of final cDNA for ligation used in traditional cDNA library construction. Thus, full length cDNA is enriched in the final PCR product for cloning.

Cloning into sequencing vectors is prone to variation. During library construction, the final ligation step is one of the most critical steps for deep coverage library generation. Traditionally, cloning is usually performed using DNA ligase. Recently, several *in vitro* recombination enzymes have been commercialized. Among those, the Cre-lox system has been used extensively for *in vivo* recombination. The directional *in vitro* recombination cloning (RC) using lambda-based recombination system is quite efficient (Hartley et al., 2000; Ohara and Temple, 2001). The original enzymes were modified to confer specificity, providing directionality in recombination. Furthermore, there is no net loss or gain of sequences that are subject to recombination. With a magazine of vector systems and modularity, it is possible to transfer library inserts in a donor vector to any destination vector with appropriate recombination sites without loss of any sequence. Moreover, the efficiency of the reaction is quite high, easily surpassing conventional DNA ligase. Thus, recombinational cloning is an attractive system for the generation of complex DNA libraries. Furthermore, the starting material only requires ~30bp sequence

for recombination which is readily amenable to PCR amplification. Here, we report an efficient method that combines the advantages of SMART and recombinational cloning (GATEWAY) for the generation of normalized cDNA libraries.

Materials and methods

Total RNA preparation

RNA was prepared from mycelia of the rice blast fungus, *Magnaporthe grisea* 70-15 strain. Fungal tissue was collected 48 hrs after growth in liquid complete medium (CM; 6 g yeast extracts, 6 g casein, 10 g sucrose, and 1 ml of *A. nidulans* trace element solution per liter), and washed with sterile distilled water. The sample was ground with liquid nitrogen to a fine powder. The sample was extracted with Trizol following the manufacturer's instruction. Poly-A RNA was isolated using PolyA tract kit from Promega.

cDNA generation and cloning

The method presented here is shown schematically in Figure 2.1. Before cDNA synthesis, the RNA sample was size fractionated using a Chromaspin DEPC-1000 column (Clontech) to enrich for high molecular weight mRNA. First strand cDNA was generated as follows: mRNA sample was mixed with 1 μ l of 10 μ M of attB2-First oligolinker (GGGGACCACTTTGTACAAGAAAGCTGGGTTTTTTTTTTTTTTTTTTTTTTT TTV) and 1 μ l attB1-Switch oligo (GGGGACAAGTTTGTACAAAAAAGCAGGCT rGrGrG), where r denotes ribonucleotide. The mix was heated at 72 °C for 2 min to

resolve secondary structure and priming site for attB2-First. The mix was quenched on ice for 5 min. Then, 1 µl of 10 mM dNTP, 1 µl of RNase inhibitor (RNasin), 2 µl of RT buffer, and 1 µl of Superscript III reverse transcriptase (Invitrogen) were added. The RT reaction mix was transferred to 42 °C, and incubated for 30 min at 42 °C, and 30 min at 55 °C. The first-strand synthesis reaction was then subjected to alkaline RNA hydrolysis by adding 1 µl of 25 mM NaOH and incubation at 68 °C for 30 min. Resulting ss-cDNA was amplified using 2 µl of the reaction mix and the primers attB1 (GGGGACAAGTTTGTACAAAAAAGCAGGCT) and attB2 (GGGGACCACTTTGTACAAGAAAGCTGGGT). We used Expand Hi-Fidelity PCR system (Roche) or other hi-fidelity Taqs in a total volume of 100 µl. PCR conditions were as follows: 94 °C for 2 min for initial denaturation followed by 22~26 cycles of 94 °C for 20 sec, 65 °C for 30 sec, and 68 °C for 5 min. PCR products were run on 0.9 % regular agarose gel for average size estimation. PCR products were purified and size fractionated using Chromaspin 1000. Cloning was done using BP clonase from Invitrogen following manufacturer's instruction. The donor vector was pDONR221 (Invitrogen). The BP reaction mix was extracted with phenol/ chloroform and precipitated with ethanol. The resulting pellet was resuspended in 5 µl of distilled water and used for transformation into electrocompetent *E. coli* DH10B cells. For comparison of this method with the existing SMART method, the same starting material was used in parallel following the manufacturer's instructions.

cDNA library normalization

cDNA library normalization was performed using double strand specific nuclease (DSN) (Zhulidov et al., 2004). Briefly, first strand cDNA was generated using the primers noted above. At the completion of the first strand cDNA synthesis, cDNA was extracted using phenol/chloroform, and subsequently precipitated using ethanol. The sample was resuspended in distilled water. Denaturation was done at 98 °C for 3 min and reassociation was done at 70 °C for 4 hrs. After 4 hrs, the sample was digested with DSN to remove reassociated DNA/RNA heteroduplexes as described. The sample was amplified using PCR for 20 cycles. The result was compared to non-normalized cDNA on a regular agarose gel.

Sequence analysis

99 clones were randomly picked for sequence analysis. Sequencing was done using M13 forward primer to verify the 5' end of the insert. Sequencing was done using BigDye terminator chemistry. The sequencing data was analyzed using Phred/Phrap (Ewing et al., 1998). Sequence assembly was done using CAP3 (Huang and Madan, 1999). BLAST analysis was done against GenBank database and in-house BLAST facilities.

Results

cDNA library construction strategy

To construct a full length cDNA library, a method that utilizes the template switching mechanism was exploited. As depicted in Figure 1, first strand cDNA was

generated using a linker with the oligo-dT and attB2 sequence. The switching oligo contained attB1 and 3 bases of ribo-Gs. To check the efficiency of the primer design, the cDNA generation procedure was compared to the original SMART cDNA library construction procedure. Figure 2.2 shows that the cDNA generated from the SMART and GATEWAY-SMART shared similar size profiles and banding patterns. This shows that cDNA generation is efficient and that the switching oligo worked correctly at the 5' end of the transcript. Also, both methods used a directional cloning system.

cDNA library normalization

cDNA library was normalized using DSN. cDNA profiles were compared between the samples before and after the normalization. As shown in Figure 2.3, cDNA bands were observed in the non-normalized library. However, in the normalized cDNA sample, no discrete bands were apparent, suggesting the sample was indeed normalized. This was further confirmed by macroarray hybridization using randomly picked clones. Hybridization with non-normalized cDNA exhibited strong spots with high signal intensities (Figure 2.4A). These clones represent highly abundant cDNA species. However, hybridization with normalized cDNA showed diminished signal intensities with more even distribution of signals (Figure 2.4B). Taken together, these data strongly indicate that the normalization procedure was effective.

Cloning

One of the critical factors in deep coverage cDNA library construction is the cloning step. Unlike the typical ligation-mediated cloning step, we found the GATEWAY

recombination to be efficient and consistent. For cloning the cDNA library, pDONR210 was used. Recombinant cloning resulted in $>10^6$ independent clones. It is well known that small fragments are preferentially cloned ligase-based system. Hartley et al suggested that the GATEWAY system showed less bias toward small fragments (Hartley et al., 2000; Ohara and Temple, 2001). Randomly picked colonies were mini-prepped and insert size was determined by restriction enzyme digestion. The recombinational sites (attB1/2 and attBL1/2) have flanking BsrGI restriction enzyme sites. The size determination by restriction enzyme is shown in the Figure 2.5. The average insert sizes for non-normalized and normalized libraries were 0.9 kb and 1.1 kb, respectively. The analysis of other clones from other independent experiments yielded no evidence for a size bias. Thus, our data suggest that the normalization procedure retained a broad spectrum of full length clones.

Sequencing

To determine the enrichment of full length cDNAs, 5' end sequencing was done using M13 forward primer. Ninety nine sequences were analyzed. Fifty sequences were from non-normalized library, and 49 sequences were from the normalized cDNA library. Assembly of the non-normalized cDNA library using CAP3 resulted in 5 contigs and 37 singletons. These contigs contained 2~4 clones and accounted for a total of 13 sequences. Thus, the non-normalized cDNA library showed 26% of redundancy. In stark contrast to the non-normalized library, no contig was formed in normalized cDNA library, again strongly suggesting that the methods we used were effective.

BLAST analysis was done using the obtained sequence data, and is summarized in Table 2.1. Similar number of hits to *M. grisea* genomic sequence and other sequences in Genbank were found for both non-normalized and normalized cDNA sequences. However, almost twice the number of hits was generated from non-normalized cDNA compared to the normalized cDNA sequences when compared to *M. grisea* ESTs. The discovery of a high number of novel EST in the normalized cDNA sequences, again, strongly indicated that normalization was effective.

In 5' region of the sequences, a putative start codon (ATG) was found in most every sequence. From the 99 sequences analyzed by BLASTX, 47 matched an annotated *M. grisea* protein sequence (Table 2.2). Among these, we were able to identify 26 ATG start codon (55% of sequences retrieved) based on the predicted annotation. Further manual inspection on the predicted ORFs suggested that at least 70% of the sequences encompass the start codon. Thus, sequencing data suggests that this SMART in combination with recombinational cloning is an efficient strategy for constructing normalized cDNA libraries.

Discussion

In this report, we generated a cDNA library using SMART and GATEWAY recombinational cloning system. The main advantage of this method is the ease of generating full length enriched cDNAs. The template switching mechanism is enhanced by the presence of a 5' cap structure and when the RT reaction is carried out at an elevated temperature (Schmidt and Mueller, 1999). These facts make SMART a system of choice for cDNA library construction. Cloning systems based on ligase usually require

restriction enzyme digestion. The PCR amplification of cDNA showed that the amplification pattern of the original SMART and recombinational SMART is almost identical. First, PCR amplification will be exponential for the full length cDNA where template switching has occurred while linear amplification will be dominant for templates without template switching. Second, cDNA with internal priming will be amplified less efficiently due to the PCR suppression effect (Matz et al., 1999; Shagin et al., 1999). Finally, use of long primers (~30 bases) should exclude non-specific amplification. In addition, cDNA with internal priming will not be cloned since the cloning is directional and requires two different sequences for recombination. Even though a large portion of sequenced cDNA clones appeared to be full length, some transcripts may not have been. There are several possible explanations. First, template switching is more efficient with full length transcripts. However, template switching does occur for truncated transcripts, thus, producing non-full length cDNAs. Quality of RNA sample is of prime importance for SMART. Second, Gene predictions are notoriously inaccurate. For example, BLAST analysis of the RADNLIB001af2_G11 clone indicated that annotation of this gene was not correctly modeled. The sequence showed a similarity to a previously characterized mitochondrial NADH-ubiquinone oxidoreductase 19.3 kDa subunit in *Neurospora crassa* (Sousa et al., 1999). However, the homologous protein MG07256.4 in *M. grisea* was not predicted correctly.

The overall strategy described in this work for generating cDNA libraries offers several advantages compared to other methods. First, this method reduces hand-on time due to the elimination of several steps that are required for cloning (for example, restriction enzyme digestion and prolonged ligation steps). Second, deeper coverage

libraries can be generated due to high cloning efficiency compared to the DNA ligase with less variation. Third, size bias toward smaller fragment is less pronounced according to other literatures (Hartley et al., 2000; Ohara and Temple, 2001). It remains to be determined whether increasing the size of vector may affect cloning of larger fragments as in ligase mediated cloning (Carninci et al., 2001).

From limited sequencing, we were able to identify novel ESTs that were not identified previously in large scale (>40,000) EST sequencing projects. The finding of the novel transcripts might be ascribed to several factors. First, the cDNA library in this work was normalized. Thus, representation of rare transcripts in the library has been achieved. Second, we employed PCR amplification of cDNA. PCR amplification might distort the proportion of cDNAs, deviating from the original proportion. This might lead to the representation of rare messages. Regardless, the data suggest that cDNA library normalization was effective, and there are many much more transcripts remaining to be discovered, particularly since we sequenced relatively few numbers of clones. Furthermore, large-scale EST sequencing analysis suggests that current gene annotation might be significantly underestimated, and the actual number of the “transcriptional units” might be larger by an order of magnitude (The FANTOM Consortium et al., 2005).

In this work, we integrated recombinational cloning procedure with SMART cDNA library construction technology. We showed that this method is amenable to library normalization step, and that normalization was effective. Thus, this method represents an efficient strategy for construction of normalized full length cDNA libraries.

References

1. Adams,M.D., Kelley,J.M., Gocayne,J.D., Dubnick,M., Polymeropoulos,M.H., Xiao,H., Merril,C.R., Wu,A., Olde,B., Moreno,R.F., and . (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252, 1651-1656.
2. Bonaldo,M.F., Lennon,G., and Soares,M.B. (1996). Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Res.* 6, 791-806.
3. Carninci,P., Shibata,Y., Hayatsu,N., Itoh,M., Shiraki,T., Hirozane,T., Watahiki,A., Shibata,K., Konno,H., Muramatsu,M., and Hayashizaki,Y. (2001). Balanced-size and long-size cloning of full-length, cap-trapped cDNAs into vectors of the novel lambda-FLC family allows enhanced gene discovery rate and functional analysis. *Genomics* 77, 79-90.
4. Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y. (1998). Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. *Proc. Natl. Acad. Sci U. S. A* 95, 520-524.
5. Das,M., Harvey,I., Chu,L.L., Sinha,M., and Pelletier,J. (2001). Full-length cDNAs: more than just reaching the ends. *Physiol Genomics* 6, 57-80.

6. Ewing,B., Hillier,L., Wendl,M.C., and Green,P. (1998). Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res.* 8, 175-185.
7. Gubler,U. and Hoffman,B.J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* 25, 263-269.
8. Hartley,J.L., Temple,G.F., and Brasch,M.A. (2000). DNA cloning using In vitro site-specific recombination. *Genome Res.* 10, 1788-1795.
9. Huang,X. and Madan,A. (1999). CAP3: A DNA sequence assembly program. *Genome Res.* 9, 868-877.
10. Lamesch,P., Milstein,S., Hao,T., Rosenberg,J., Li,N., Sequerra,R., Bosak,S., Doucette-Stamm,L., Vandenhaute,J., Hill,D.E., and Vidal,M. (2004). *C. elegans* ORFeome version 3.1: increasing the coverage of ORFeome resources with improved gene predictions. *Genome Res.* 14, 2064-2069.
11. Luan,C.H., Qiu,S., Finley,J.B., Carson,M., Gray,R.J., Huang,W., Johnson,D., Tsao,J., Reboul,J., Vaglio,P., Hill,D.E., Vidal,M., Delucas,L.J., and Luo,M. (2004). High-throughput expression of *C. elegans* proteins. *Genome Res.* 14, 2102-2110.
12. Matz,M., Shagin,D., Bogdanova,E., Britanova,O., Lukyanov,S., Diatchenko,L., and Chenchik,A. (1999). Amplification of cDNA ends based on template-switching effect and step-out PCR. *Nucleic Acids Res.* 27, 1558-1560.

13. Ohara,O. and Temple,G. (2001). Directional cDNA library construction assisted by the in vitro recombination reaction. *Nucleic Acids Res.* 29, E22.
14. Reboul,J., Vaglio,P., Rual,J.F., Lamesch,P., Martinez,M., Armstrong,C.M., Li,S., Jacotot,L., Bertin,N., Janky,R., Moore,T., Hudson,J.R., Jr., Hartley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S., Endress,G.A., Jenna,S., Chevet,E., Papatotiropoulos,V., Tolias,P.P., Ptacek,J., Snyder,M., Huang,R., Chance,M.R., Lee,H., Doucette-Stamm,L., Hill,D.E., and Vidal,M. (2003). *C. elegans* ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression. *Nat Genet.* 34, 35-41.
15. Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S., Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O., Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T., Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C., Vandenhaute,J., Cusick,M.E., Albala,J.S., Hill,D.E., and Vidal,M. (2004). Human ORFeome version 1.1: a platform for reverse proteomics. *Genome Res.* 14, 2128-2135.
16. Rual,J.F., Venkatesan,K., Hao,T., Hirozane-Kishikawa,T., Dricot,A., Li,N., Berriz,G.F., Gibbons,F.D., Dreze,M., Ayivi-Guedehoussou,N., Klitgord,N., Simon,C., Boxem,M., Milstein,S., Rosenberg,J., Goldberg,D.S., Zhang,L.V., Wong,S.L., Franklin,G., Li,S., Albala,J.S., Lim,J., Fraughton,C., Llamosas,E., Cevik,S., Bex,C., Lamesch,P., Sikorski,R.S., Vandenhaute,J., Zoghbi,H.Y., Smolyar,A., Bosak,S., Sequerra,R., Doucette-Stamm,L., Cusick,M.E., Hill,D.E.,

- Roth,F.P., and Vidal,M. (2005). Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437, 1173-1178.
17. Schmidt,W.M. and Mueller,M.W. (1999). CapSelect: a highly sensitive method for 5' CAP-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. *Nucl. Acids Res.* 27, e31.
 18. Shagin,D.A., Lukyanov,K.A., Vagner,L.L., and Matz,M.V. (1999). Regulation of average length of complex PCR product. *Nucleic Acids Res.* 27, e23.
 19. Soares,M.B., Bonaldo,M.F., Jelene,P., Su,L., Lawton,L., and Efstratiadis,A. (1994). Construction and characterization of a normalized cDNA library. *Proc. Natl Acad. Sci U. S. A* 91, 9228-9232.
 20. Sousa,R., Barquera,B., Duarte,M., Finel,M., and Videira,A. (1999). Characterisation of the last Fe-S cluster-binding subunit of *Neurospora crassa* complex I. *Biochim Biophys Acta.* 1411, 142-146.
 21. The FANTOM Consortium, Carninci,P., Kasukawa,T., Katayama,S., Gough,J., Frith,M.C., Maeda,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C., Kodzius,R., Shimokawa,K., Bajic,V.B., Brenner,S.E., Batalov,S., Forrest,A.R.R., Zavolan,M., Davis,M.J., Wilming,L.G., Aidinis,V., Allen,J.E., Ambesi-Impiombato,A., Apweiler,R., Aturaliya,R.N., Bailey,T.L., Bansal,M., Baxter,L., Beisel,K.W., Bersano,T., Bono,H., Chalk,A.M., Chiu,K.P., Choudhary,V., Christoffels,A.,

Clutterbuck,D.R., Crowe,M.L., Dalla,E., Dalrymple,B.P., de Bono,B., Gatta,G.D.,
di Bernardo,D., Down,T., Engstrom,P., Fagiolini,M., Faulkner,G., Fletcher,C.F.,
Fukushima,T., Furuno,M., Futaki,S., Gariboldi,M., Georgii-Hemming,P.,
Gingeras,T.R., Gojobori,T., Green,R.E., Gustincich,S., Harbers,M., Hayashi,Y.,
Hensch,T.K., Hirokawa,N., Hill,D., Huminiecki,L., Iacono,M., Ikeo,K., Iwama,A.,
Ishikawa,T., Jakt,M., Kanapin,A., Katoh,M., Kawasawa,Y., Kelso,J., Kitamura,H.,
Kitano,H., Kollias,G., Krishnan,S.P.T., Kruger,A., Kummerfeld,S.K.,
Kurochkin,I.V., Lareau,L.F., Lazarevic,D., Lipovich,L., Liu,J., Liuni,S.,
McWilliam,S., Babu,M.M., Madera,M., Marchionni,L., Matsuda,H., Matsuzawa,S.,
Miki,H., Mignone,F., Miyake,S., Morris,K., Mottagui-Tabar,S., Mulder,N.,
Nakano,N., Nakauchi,H., Ng,P., Nilsson,R., Nishiguchi,S., Nishikawa,S., Nori,F.,
Ohara,O., Okazaki,Y., Orlando,V., Pang,K.C., Pavan,W.J., Pavesi,G., Pesole,G.,
Petrovsky,N., Piazza,S., Reed,J., Reid,J.F., Ring,B.Z., Ringwald,M., Rost,B.,
Ruan,Y., Salzberg,S.L., Sandelin,A., Schneider,C., Schonbach,C., Sekiguchi,K.,
Semple,C.A.M., Seno,S., Sessa,L., Sheng,Y., Shibata,Y., Shimada,H., Shimada,K.,
Silva,D., Sinclair,B., Sperling,S., Stupka,E., Sugiura,K., Sultana,R., Takenaka,Y.,
Taki,K., Tammoja,K., Tan,S.L., Tang,S., Taylor,M.S., Tegner,J., Teichmann,S.A.,
Ueda,H.R., van Nimwegen,E., Verardo,R., Wei,C.L., Yagi,K., Yamanishi,H.,
Zabarovsky,E., Zhu,S., Zimmer,A., Hide,W., Bult,C., Grimmond,S.M.,
Teasdale,R.D., Liu,E.T., Brusica,V., Quackenbush,J., Wahlestedt,C., Mattick,J.S.,
Hume,D.A., RIKEN Genome Exploration Research Group and Genome Science
Group (Genome Network Project Core Group), Kai,C., Sasaki,D., Tomaru,Y.,
Fukuda,S., Kanamori-Katayama,M., Suzuki,M., Aoki,J., Arakawa,T., Iida,J.,

- Imamura,K., Itoh,M., Kato,T., Kawaji,H., Kawagashira,N., Kawashima,T., Kojima,M., Kondo,S., Konno,H., Nakano,K., Ninomiya,N., Nishio,T., Okada,M., Plessy,C., Shibata,K., Shiraki,T., Suzuki,S., Tagami,M., Waki,K., Watahiki,A., Okamura-Oho,Y., Suzuki,H., Kawai,J., and Hayashizaki,Y. (2005). The Transcriptional landscape of the mammalian genome. *Science* 309, 1559-1563.
22. Zhu,Y.Y., Machleder,E.M., Chenchik,A., Li,R., and Siebert,P.D. (2001). Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *Biotechniques* 30, 892-897.
23. Zhulidov,P.A., Bogdanova,E.A., Shcheglov,A.S., Vagner,L.L., Khaspekov,G.L., Kozhemyako,V.B., Matz,M.V., Meleshkevitch,E., Moroz,L.L., Lukyanov,S.A., and Shagin,D.A. (2004). Simple cDNA normalization using kamchatka crab duplex-specific nuclease. *Nucleic Acids Res.* 32, e37.

Table 2.1 BLAST analysis of non-normalized and normalized cDNA libraries

BLAST/Databases ^a	Non-normalized (% hit) ^b	Normalized (% hit)
BLASTN/MG_genome	43 (86)	46 (94)
BLASTX/nr	22 (44)	25 (51)
BLASTN/MG_EST	36 (72)	15 (31)

^a BLAST programs and BLAST databases used for analysis. MG_genome, nr, and MG_EST are BLAST databases that contain *M. grisea* genomic sequences, non-redundant GenBank sequences, and *M. grisea* ESTs, respectively.

^b The number of matches found with e-value cutoff of $1e^{-5}$. % hit denotes the percent of non-normalized and normalized sequences with a significant match in their respective databases.

Total of 99 sequences were used for BLAST analysis (50 from non-normalized and 49 from normalized cDNA libraries, respectively). Comparison of hits generated by BLAST against genomic and EST sequences suggests that less hits were generated from EST sequences. This suggests that normalization was effective for recovering rare messages.

Table 2.2 BLAST analysis of cDNA sequences

Clone ^a	FL ^b	Genome ^c	Genes
RLIB001af2_E04	F	<i>Mg</i>	gb EAA47404.1 hypothetical protein MG02647.4
RLIB001af2_A05	F	<i>Mg</i>	gb EAA53444.1 hypothetical protein MG07721.4
NLIB001af2_G12	F	<i>Mg</i>	gb EAA54132.1 hypothetical protein MG02117.4
NLIB001af2_G11	F	<i>Nc</i>	sp O47950 NUKM_NEUCR NADH-ubiquinone oxidoreductase
NLIB001af2_G06	F	<i>Mg</i>	gb EAA47467.1 hypothetical protein MG02710.4
NLIB001af2_G05	F	<i>Mg</i>	gb EAA52611.1 hypothetical protein MG05303.4
NLIB001af2_G04	F	<i>Mg</i>	gb EAA53813.1 hypothetical protein MG09563.4
NLIB001af2_F07	F	<i>Mg</i>	gb EAA54132.1 hypothetical protein MG02117.4
NLIB001af2_F03	F	<i>Mg</i>	gb EAA54206.1 predicted protein
NLIB001af2_E11	F	<i>Mg</i>	gb EAA51733.1 predicted protein ref XP_360785.1
NLIB001af2_E09	F	<i>Mg</i>	gb EAA56693.1 hypothetical protein MG07048.4
NLIB001af2_E08	F	<i>Mg</i>	gb EAA48691.1 hypothetical protein MG00349.4
NLIB001af2_D10	F	<i>Mg</i>	gb EAA56011.1 hypothetical protein MG01662.4
NLIB001af2_D08	F	<i>Mg</i>	gb EAA51565.1 hypothetical protein MG03160.4
NLIB001af2_D05	F	<i>Mg</i>	gb EAA52949.1 hypothetical protein MG06077.4
NLIB001af2_D04	F	<i>Mg</i>	gb EAA54688.1 hypothetical protein MG05480.4
NLIB001af2_C10	F	<i>Mg</i>	gb EAA54844.1 hypothetical protein MG05635.4
NLIB001af2_B10	F	<i>Mg</i>	gb EAA55090.1 hypothetical protein MG06747.4
NLIB001af2_B07	F	<i>Mg</i>	gb AAW69347.1 ATP synthase oligomycin sensitivity
NLIB001af2_A12	F	<i>Mg</i>	gb EAA47543.1 hypothetical protein MG02786.4
jreg001xg15.b2	F	<i>Mg</i>	gb EAA56278.1 hypothetical protein MG06249.4
jreg001xg07.b2	F	<i>Gz</i>	gb EAA67793.1 CALM_NEUCR Calmodulin (CaM)
jreg001xg01.b2	F	<i>Mg</i>	gb EAA56000.1 hypothetical protein MG01651.4
jreg001xe13.b2	F	<i>Mg</i>	gb EAA51541.1 hypothetical protein MG03136.4
jreg001xc21.b2	F	<i>Mg</i>	gb EAA50345.1 hypothetical protein MG04104.4
jreg001xc09.b2	F	<i>Mg</i>	gb EAA48946.1 hypothetical protein MG00604.4
jreg001xe15.b2	F	<i>Mg</i>	gb EAA56278.1 hypothetical protein MG06249.4
jreg001xg21.b2	F	<i>An</i>	gb EAA58337.1 hypothetical protein AN5828.2
NLIB001af2_A11	F	<i>Nc</i>	ref XP_331829.1 hypothetical protein gb EAA34819.1
NLIB001af2_H04	?	<i>Mg</i>	gb EAA56783.1 hypothetical protein MG07138.4
jreg001xc03.b2	?	<i>Mg</i>	hypothetical protein MG11035.4
NLIB001af2_C08	?	<i>Mg</i>	gb EAA47369.1 hypothetical protein MG02612.4
jreg001xg01.b2	?	<i>Nc</i>	ref XP_331643.1 predicted protein
NLIB001af2_C12	?	<i>Mg</i>	gb EAA53444.1 hypothetical protein MG07721.4
jreg001xc11.b2	?	<i>Mg</i>	gb EAA53963.1 hypothetical protein MG01948.4
RLIB001af2_D04	?	<i>Nc</i>	ref XP_322915.1 predicted protein gb EAA32104.1
jreg001xe07.b2	?	<i>Mg</i>	gb EAA50003.1 hypothetical protein MG03762.4
jreg001xe01.b2	?	<i>Mg</i>	gb EAA48461.1 hypothetical protein MG00119.4
jreg001xg03.b2	?	<i>Mg</i>	gb EAA57279.1 hypothetical protein MG08248.4
jreg001xe07.b2	?	<i>Mg</i>	gb EAA57279.1 hypothetical protein MG08248.4
NLIB001af2_E04	NF	<i>Mg</i>	gb EAA49307.1 hypothetical protein MG00965.4
NLIB001af2_F05	NF	<i>Mg</i>	gb EAA50253.1 hypothetical protein MG04012.4
NLIB001af2_E12	NF	<i>Gz</i>	gb EAA70180.1 hypothetical protein FG00101.1
RLIB001af2_F03	NF	<i>Nc</i>	ref XP_326034.1 hypothetical protein gb EAA33755.1
RLIB001af2_G04	NF	<i>Mg</i>	gb EAA51376.1 hypothetical protein MG09393.4
RLIB001af2_D03	NF	<i>Mg</i>	gb EAA51376.1 hypothetical protein MG09393.4
RLIB001af2_B03	NF	<i>Mg</i>	gb EAA50413.1 hypothetical protein MG04172.4

Table 2.2 (continued)

^a Sequences with RLIB- and jreg- suffixes correspond to clones from the non-normalized cDNA library, and NLIB sequences are from the normalized library.

^b F denotes a confirmed full length cDNA, ? corresponds to sequences that support the evidence of full length or near full length, and NF is non-full length determined by BLAST hits. Hits are genomes to which the sequences matched.

^c Genome from which matching sequence was found. An: *Aspergillus nidulans*, Mg: *M. grisea*, Nc: *Neurospora crassa*, Gz: *Gibberella zeae*.

Figure legends

Figure 2.1 cDNA synthesis scheme.

First strand cDNA is synthesized using attB2-T primer to anchor 3' end of transcripts. During the reverse transcription reaction, template switching occurs at the 5' end of the transcript. The template switching defines the 3' end of the cDNA. The resulting cDNA is subject to PCR amplification using the LA-PCR procedure.

Figure 2.2 Comparison of cDNA generation.

cDNA was generated using original SMART cDNA library construction method and recombinational SMART cDNA procedure described in methods. The first strand synthesis mix was PCR amplified using SMART oligos and attB1 and attB2 primers. Lane M is a size marker (Sigma D7085). Lanes 1 and 2 are original SMART and recombinational SMART, respectively.

Figure 2.3 cDNA library normalization.

cDNA library was normalized using double strand specific nuclease (DSN) as described in the materials and methods section. Lane M: Size marker (Sigma D7085), Lane S: cDNA library before normalization, Lane N: cDNA library after normalization.

Figure 2.4 cDNA filter hybridization of normalized cDNA library.

A high density cDNA filter was prepared, and hybridization was done with non-normalized and normalized cDNAs as probes. Compared to the non-normalized cDNA (A), the normalized cDNA probe (B) showed great uniformity of signal intensities.

Figure 2.5 The size distribution of the non-normalized and normalized cDNA libraries.

The cDNA clones were digested with BsrGI. The restriction enzyme cuts attB1 and attB2 sequences, freeing the insert. Panel (A) and (B) is the BsrGI digestion of the clones from the non-normalized and normalized cDNA libraries, respectively. The size distribution of clones in each library showed little apparent differences. M indicates the size marker lane.

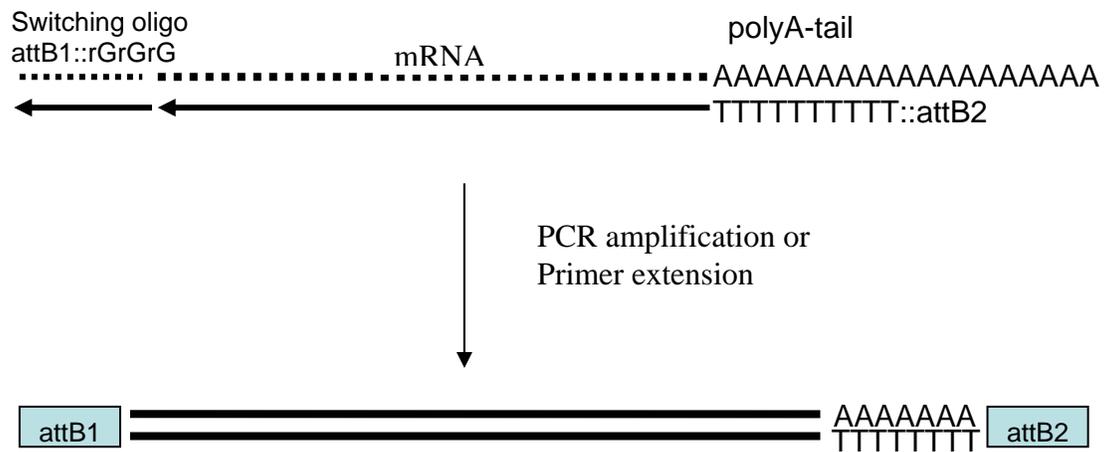


Figure 2.1 cDNA synthesis scheme

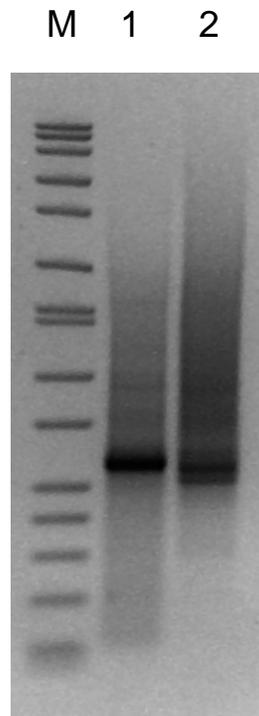


Figure 2.2 Comparison of cDNA generation

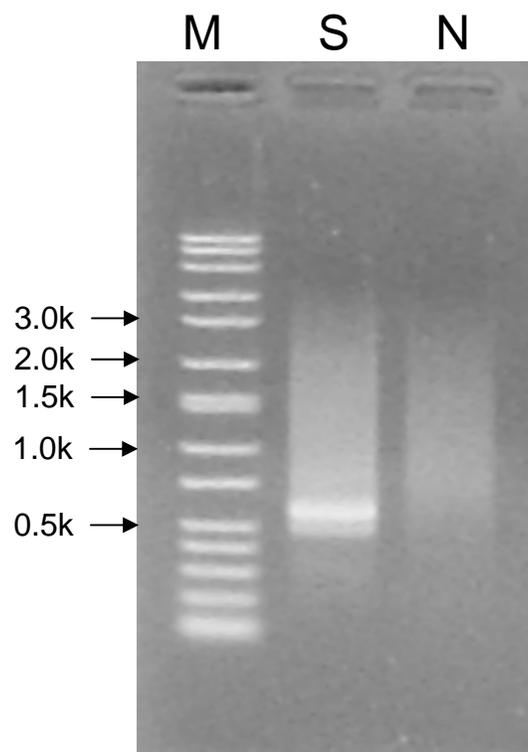
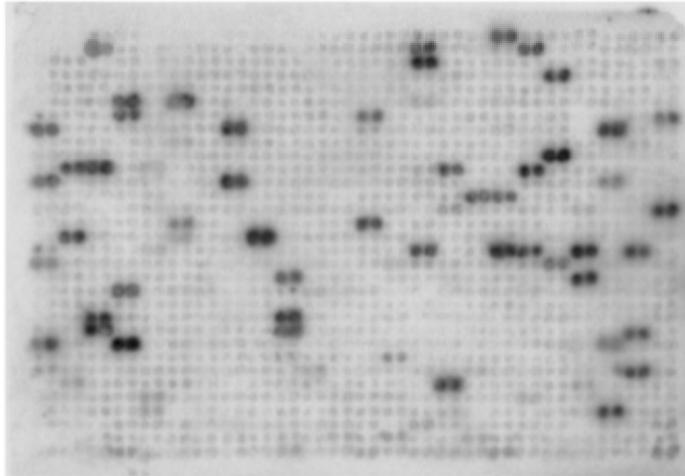


Figure 2.3 cDNA library normalization

(A)



(B)

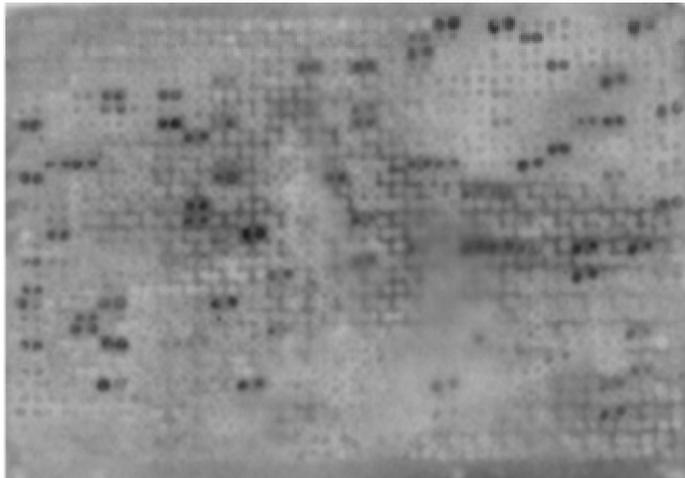


Figure 2.4 cDNA filter hybridization of normalized cDNA library

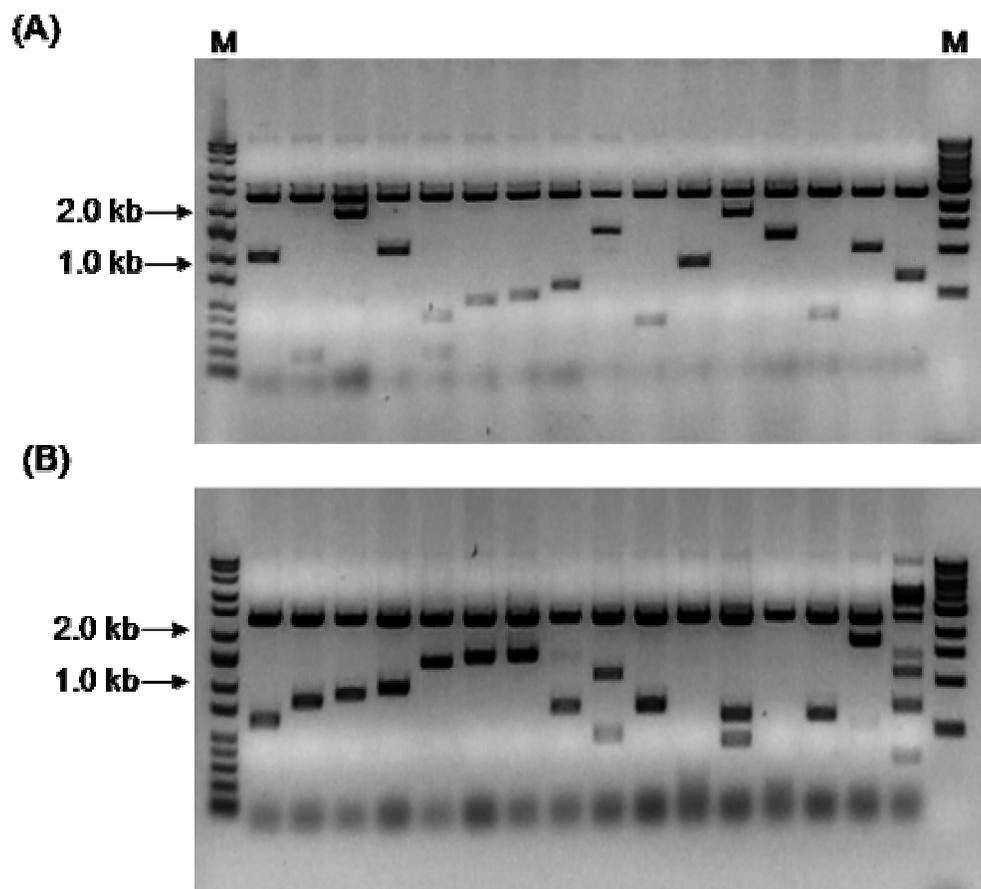


Figure 2.5 The size distribution of the non-normalized and normalized cDNA libraries

Chapter Three

Genome-wide microarray analysis of secretome in the rice blast fungus,

Magnaporthe grisea

Jun Seop Jeong and Ralph A. Dean

Center for Integrated Fungal Research (CIFR), North Carolina State University, Raleigh,
NC 27695

Abstract

Proteins that are represented extracellularly such as secreted and membrane-bound proteins play important roles during the interaction of plant pathogenic fungi with their hosts. Extracellular, including membrane-bound, proteins are targeted to the ER and secretory pathways during transcription and translation. Whole genome microarrays can be used as a functional tool to investigate this targeting. In this work, we performed subcellular fractionation to isolate microsome-associated mRNA and free cytosolic mRNA followed by microarray analysis. Also, predicted open reading frames were analyzed for the presence of sequences that potentially affect secretion and membrane-association such as signal peptide, transmembrane domains, and GPI-anchor signals. Microarray analysis suggests that nascent proteins appear to be targeted to secretory pathway either co-translationally or post-translationally. The hydrophobicity of the signal peptide did not appear to be the primary determinant for routing. Fractionation was affected by factors such as presence of a signal peptide and presence of transmembrane domains as expected. Importantly, dynamic localization of transcripts between microsomal and free cytosolic fraction was noticed in different growth stages. Also, retrotransposon elements were shown to be enriched in the microsomal fraction, suggesting the endomembrane system may be an intracellular habitat for retroviruses. The data will be useful to identify factors that determine routing of proteins through the secretory machinery.

Introduction

The growth and morphological development of fungi is the outcome of continuous communication with the environment. Availability of nutrients affects mycelial ramification and developmental programs such as sporulation and mating. At the interface of the communication are membrane sensors and secreted proteins. Those proteins play important roles in sensing ambient conditions and acquisition of nutrients. Proteins that are found in the extracellular space have been the subject of extensive research in plant-pathogen interactions. In the rice blast fungus, *Magnaporthe grisea*, secreted proteins such as a hydrophobin, MPG1, and CBP, a chitin-binding protein, were shown to be critical factors for infection-related development (Kamakura et al., 2002; Talbot et al., 1996). Also, integral membrane proteins, such as PTH11 putative surface sensor (DeZwaan et al., 1999), and PLS1, a tetraspanin-like protein, were shown to be required for the differentiation of the penetration peg during infection (Clergeot et al., 2001).

The genome of the rice blast fungus has been sequenced recently (Dean et al., 2005). Analysis of the genome revealed over 11,000 genes in the 40 Mbp genome. The genome is equipped with a battery of secreted proteins that may play important roles during the life cycle and infection process. Unlike its close relative, the saprophyte *Neurospora crassa*, there appears to be more genes content that are important for signal perception and for interaction with the environment. These include putative G-protein coupled receptors (GPCR) possessing the CFEM domain and proteins with novel chitin binding motifs. Interestingly, no cutinase was found in *N. crassa* genome while multiple

cutinases were found in the rice blast fungus, suggesting niche-specific gene expansion (Galagan et al., 2003). The genome of the rice blast fungus has a repertoire of repetitive elements. The repetitive elements such as retrotransposons or transposons can create variation in the genome and phenotypes by mobilizing and inserting into new loci, acting as an evolutionary engine creating new phenotypes (Fudal et al., 2005; Kang et al., 2001). As documented in many pathogens, proteins targeted to the secretory pathway such as membrane proteins, secreted proteins, and organellar proteins play critical roles in each step of infection and colonization. For *M. grisea*, the infection process begins with the attachment of spores to the host surface. Attachment is mediated by a spore tip mucilage (STM) which provides a tenacious binding to the surface (Hamer et al., 1988), and is composed of glycoproteins. Shortly after attachment, germination follows. One of the crucial features of *M. grisea* that appears during the infection process is the formation of the appressorium (Dean, 1997). In contrast to the filamentous form of vegetative growth, the appressorium is a dome-like structure, which becomes heavily melanized towards maturation. Detailed microscopic analysis has revealed that the appressorium is enclosed in a melanin layer with the exception of a region that is directly contacting the plant surface. Within the appressorium, the concentration of glycerol builds up to generate turgor pressure. The turgor is channeled so that it helps propel a penetration peg through the host surface (Howard et al., 1991). Thus, the rice blast fungus does not appear to rely on hydrolytic enzymes for penetration. However, hydrolytic enzymes might be important factors in the penetration process (Dean et al., 2005). Gene expression profiling suggests that cutinase(s) are induced during appressorium formation. Also, appressorium formation can be induced at a nanomolar level of cutin monomers on non-inductive

surfaces (Gilbert et al., 1996). It has been recently shown that the rice blast fungus can undergo an alternative developmental process used by soil-borne plant pathogenic fungi to infect through the root surface (Sesma and Osbourn, 2004). Appressorium formation does not appear to be a prerequisite for infection on root tissue, and the pathogen adopts a strategy exploited by root pathogens. However, the compatibility conferred by gene-for-gene interaction is still operative during root infection. Secreted proteins during pathogenesis often encode hydrolytic enzymes. So far, no critical role has been found in this class of secreted proteins in the rice blast fungus during pathogenic interaction (Sweigard et al., 1992; Wu et al., 1997). The effect of individual enzymes during the pathogenesis in pathogenic fungi may not be readily apparent (Schafer, 1994; Walton, 1994). Functional redundancy may compensate such as the presence of multiple cutinases in the genome (Dean et al., 2005). On the other hand, developmental roles for secreted protein during pathogenic development have been reported such as MPG1, a hydrophobin gene, and a chitin binding protein, CBP1, as mentioned earlier (Kamakura et al., 2002; Talbot et al., 1996). Cell wall degrading enzymes produced by fungi during pathogenic interaction may serve as virulence factors rather than represent determinants of pathogenicity (Schafer, 1994). However, many avirulence gene products are secreted proteins that act as pathogenic determinants in a gene-for-gene manner (Lauge and de Wit, 1998). Thus, secreted proteins and membrane-associated proteins play important roles during the infection process.

The DNA microarray is a central platform for functional genomics. Microarrays have been used mainly for the comparison of gene expression profiling from different conditions. This approach not only provides the information of genes that are

differentially regulated but also provides a diagnostic tool for the classification of samples. However, the level of transcription does not always correlate with the level of proteins (Gygi et al., 1999). Multi-step regulation of protein levels such as transcriptional, post transcriptional, translational, and post translational levels ensures that proteins are expressed at the level that are required by the cell at a given stage. To this end, translational profiling has been proposed (Pradet-Balade et al., 2001). Subcellular fractionation and target specific enrichment of transcripts provide information on the cellular localization of molecules in high throughput screening (Kopczynski et al., 1998). Separation of polysomal fraction and free fraction has been used to identify transcripts that are associated with ribosomes and are actively translated through microarray experiments (Mikulits et al., 2000). Diehn et al. showed that comparison of membrane-bound mRNA (MBP) and free cytosolic fraction (FP) using the microarray format is a complementary method to identify proteins that are secreted and/or membrane-bound (Diehn et al., 2000). A similar approach taken in a mammalian system provided information on identifying secreted proteins and correct annotation of previously known genes (Toyoda et al., 2003). Subcellular fractionated samples help provide a clue to the localization of translation protein products and functional information to the gene annotation. Thus, we took the approach of subcellular fractionation and microarray analysis to identify secreted and membrane-bound proteins in the rice blast fungus.

Samples for microarray experiment were prepared from mycelial samples grown in synthetic minimal medium and from appressoria samples induced on an artificial substratum. We found that microsomal fractions isolated from two different stages were rich in transcripts encoding proteins that contain signal peptides that target to the

extracellular space or plasma membrane (permeases and transporters) and subcellular organelles such as ER, Golgi apparatus, and vacuole. Our data suggests that there are two translocation pathways for secretory proteins: Co-translational and post-translational translocation into endoplasmic reticulum (ER). Dynamic association of transcripts to microsomal fraction was observed under different conditions, suggesting regulation of translation and localization. Also, enrichment is affected by protein structures such as transmembrane domains and lipid association sequences.

Materials and methods

Strains and culture

Magnaporthe grisea 70-15 strain was used in all experiments, for which the genome sequence is available (Dean et al., 2005). Strain 70-15 was grown on Oatmeal agar (OMA) for routine propagation and sporulation. For liquid cultures, a conidial suspension was inoculated in complete medium (CM; Yeast extracts: 6 g, Sucrose: 20 g, Casein: 6 g, thiamine 1 mg, biotin 5 µg, and 1ml/liter of trace element solution) and grown for 24 hrs. The culture was collected and washed extensively with sterile distilled water and the mycelial mass was transferred to liquid minimal medium (MM; Sucrose 10 g, *A. nidulans* trace element solution 1 ml, 20x nitrate salts 50 mL, thiamine 1 mg, and biotin 5 µg per liter). The trace element solution is NaNO₃ 60 g, KCl 5.2 g, MgSO₄7H₂O 5.2 g, KH₂PO₄ 15.2 g in 500 ml, and growth was continued for an additional 48 hrs.

RNA sample preparation

RNA samples for microarray experiments were prepared according to the published procedure with a slight modification (Stoltenburg et al., 1995). The overall scheme is depicted in Figure 3.1. The culture, grown in MM liquid medium for 48 hrs (MM48), was collected using vacuum filtration. The mycelial mass was ground with liquid nitrogen, and resuspended in extraction buffer amended with ribonucleoside vanadyl complex (10 mM final concentration) and cycloheximide (0.1 mg/ml). To remove mycelial debris and particulates, low speed centrifugation was done at 7000 rpm in Sorvall SS-34 rotor (~6000X g) to obtain a post-nuclear supernatant. The post-nuclear supernatant was used for differential centrifugation to obtain the microsomal fraction (MBP) and cytoplasmic fraction (FP) according to procedure described in the reference above.

For appressorium induction, the spore suspension was spread on cellophane film and incubated at room temperature. The formation of appressorium was monitored and the sample was prepared when approximately 30 % of germinating conidia formed appressoria. The cellophane film was ground with liquid nitrogen and resuspended in the extraction buffer. To remove the cellophane film and cell debris, the resuspended sample was passed through sterile Miracloth and subjected to differential centrifugation in the same way as the MM48 sample. To verify results obtained from differential centrifugation, sucrose gradient ultracentrifugation was performed for MM48 sample (Mechler, 1987). Comparison between differential centrifugation prep and sucrose gradient prep was done using RT-PCR analysis using selected primers.

Polysomal RNA preparation was done following the protocol employed in budding yeast with minor modifications (Kuhn et al., 2001). Briefly, ribonucleoside vanadyl complex was used as RNase inhibitor (10 mM final concentration). Cell lysate was cleared by centrifugation at 7000 rpm in Sorval SS-34 rotor (~6000X g) for 5 min. Supernatant was taken for ultracentrifugation in FFT swinging bucket rotor. Following centrifugation, the gradient was collected and O.D. monitored at 260 nm using a Duo-Flo HPLC. Fractions containing monosomes and polysomes were collected.

Microarray analysis and validation

MBP and FP sample labeling and array processing was done according to the Agilent manual. Microarray scanning was done by using Agilent G2565B DNA Microarray Scanner and Agilent G2566A Feature Extraction Software (version 6.1). The microarray data were filtered to exclude features with population outliers and nonuniformity feature outliers. Further analysis was performed for genes with MG numbers. Background intensities were carefully inspected, and the variation was low with mean intensities lower than 50 and standard deviation below 20. Thus, signal intensities above 100 were used for further analysis. Microarray experiments were performed with one technical replicate and repeated using freshly prepared samples. Detailed analyses were done using GeneSpring software (Agilent). Statistical analysis was done using the JMP package (SAS institute).

To validate the microarray data, semi-quantitative RT-PCR analysis was done as follows: RT-PCR analysis was performed using fresh samples prepared by differential centrifugation and sucrose gradients. Approximately, 2~3 µg of RNA samples were used

for reverse transcription reactions. Specific primers were purchased from IDT DNA technology. The primer sequences are listed in the Table 3.1. Twenty four to twenty six PCR cycles were performed followed by gel electrophoresis. RNA samples for semi-quantitative RT-PCR were independently prepared from the samples used for microarray. Signal peptide analysis and transmembrane domain analysis was done using SignalP (ver.3) (Bendtsen et al., 2004) and TMHMM (ver.2) (Krogh et al., 2001). For the generation of hydropathy plots, pepwindow software in the EMBOSS package was used (Rice et al., 2000). GPI-anchored protein prediction was done using Big-Pi (Eisenhaber et al., 2004) and GPI-SOM (Fankhauser and Maser, 2005a).

Immunofluorescence microscopy

Appressoria were induced for 24 hrs on cellophane film at room temperature. Chitosan-specific antibody was a kind gift from Holger Deising. The staining procedure used was as described in the literature (El Gueddari et al., 2002). Secondary antibody was FITC-conjugated anti-rabbit antibody raised in goats (Sigma). The staining was done at 1:200-dilution. Chitin in appressoria and penetration pegs was stained with wheat germ agglutinin (WGA) conjugated with TRITC (Sigma) in 2 µg/ml in PBS along with the chitosan staining. Microscopic observation was made using FITC and Texas Red filter set for FITC and TRITC, respectively.

Results

To identify genes that encode secreted and membrane-bound proteins in *M. grisea*, subcellular fractionation and whole genome microarray analyses were performed. Since

secretory proteins are routed to the ER during translation, transcripts that are translated on ER-bound polysomes encode proteins that are targeted to the secretory pathway. Two conditions were analyzed: a mycelial sample grown in synthetic liquid minimal medium and appressoria collected from cellophane film. We hypothesized that the mycelial sample grown in minimal media (MM48) should represent a reduced set of secreted proteins into medium since the medium contains no complex molecules that need to be digested before uptake for nutrition. On the other hand, since the appressorium (APP) is formed for penetration and infection of leaf tissues, we expected that under this condition a number of secreted/membrane-associated proteins, potentially involved in host recognition, attachment, germination, and penetration might be identified.

Analysis of microarray data

The filtering of microarray data as described in the materials and methods resulted in 7836 features for MM48 sample and 5740 features for APP sample. The top 25 transcripts that were enriched in MBP and FP fraction from each condition are shown in Table 3.2 and 3.3. The list of genes in the MBP-enriched transcripts included genes encoding proteins that are known to be targeted to the secretory pathway. Genes in the MBP fraction include many organellar proteins such as binding protein (BiP), vacuolar peptidases, protein disulfide isomerase (PDI), and calnexin. BiP and vacuolar peptidases were found to be enriched in APP and MM48 samples. However, comparison of ranked lists suggested that vacuolar proteases are induced during the appressorium formation. Calnexin and different PDIs were shown to be enriched in APP MBP, suggesting an increased protein secretion. Also, several of transporters and permeases were included.

In *Bipolaris oryzae*, UV-1 was shown to be expressed upon UV irradiation. The protein contains a putative signal peptide. Previous work suggested that this gene is expressed during appressorium formation in *M. grisea* (Kihara et al., 2001). We found the expression of the gene in the appressorial sample enriched in APP MBP, but no detectable expression was seen in MM48 samples. A versicolorin B synthase homolog was also enriched in MBP fraction in APP. The gene encodes a key enzyme in aflatoxin B1 synthesis. In *Aspergillus parasiticus*, the enzyme is localized in peri-nuclear ER, and the expression was shown to be consistent with aflatoxin production (Chiou et al., 2004). The metabolism of trehalose in the rice blast fungus was found to be important for pathogenicity (Foster et al., 2003). Two trehalase activities were reported. In our microarray data, a trehalase TRE1 (AZM11513) was found to be enriched in MBP fraction of APP, again suggesting its induction during appressorium formation. TRE1 in the rice blast fungus is known to be located in the cell wall. In contrast to the MBP fraction, the FP fraction was rich in protein translation machineries such as ribosomal proteins and translational apparatus. Also, the list included proteins that play a role in protein modification such as ubiquitination and lipid biosynthesis. However, most of the transcripts were of unknown function. Thus, the data generated by microarray is consistent with other experimental verified localization studies.

Ergosterol biosynthesis involves interplay between cytoplasmic and ER-associated enzymes components. To further explore this relationship, we investigated the partitioning of the enzymes in the ergosterol biosynthesis pathway. Early steps in ergosterol synthesis consist of mevalonate and isopentenyl pathways that reside in the cytoplasm. However, later steps of ergosterol synthesis are present in the ER and ER

associated lipid bodies. As shown in Figure 3.2, transcripts that are ER-associated enzymes function in later steps of ergosterol biosynthesis were enriched in MBP while transcripts for mevalonate and isopentenyl pathway were enriched in FP, indicating differential level of partitioning.

Computational prediction of proteins with signal peptide and membrane-bound proteins

We used protein sequence data to investigate the correlation between the microarray results and presence of signal peptide and transmembrane domains. We used MG protein sequence version 2, which has 11108 protein sequence entries. SignalP analysis was done using both Neural Network Model (NNM) and Hidden Markov Model (HMM) prediction algorithms. Among the protein sequences, NNM predicted 2024, and HMM predicted 2213 protein sequences with a signal peptide; 1832 were shared by both algorithms. Analysis of transmembrane domains using TMHMM revealed that 2011 out of 11108 proteins possessed a transmembrane domain. Comparing the data sets obtained using SignalP and TMHMM indicated that 1354 sequences lacked a transmembrane domain, but contained a signal peptide sequence, representing the potential secreted protein set. We further investigated the presence of transmembrane domains in a previously published secreted protein set in *M. grisea* (Dean et al., 2005). The published secretome set consists of 739 protein sequences. TMHMM analysis suggested that 109 sequences of this set possess a potential transmembrane domain, reducing the secretome to 630 proteins.

To determine genes that were predominantly associated with the microsomal fraction (MBP) compared to the free cytosolic fraction (FP), the microarray data were

analyzed using gene annotations and computational analysis. Expression values for genes predicted to contain a signal peptide and/or a transmembrane domain have been used to calibrate microarray data. In yeast and human cells, calibration methods such as moving average and arbitrary cutoff values using known genes have been used (Diehn et al., 2000; Toyoda et al., 2003). For our analysis, we compared the MBP:FP signal between genes with and without a predicted signal peptide or transmembrane domain.

Statistical analysis revealed that not only the signal peptide but also the transmembrane domain affected localization of transcripts (Table 3.4). The t-test (assuming non-equal variance) showed that the mean signal intensities were significantly larger for sequences possessing a signal peptide or transmembrane domain compared to genes that lack these features. Non-parametric median test gave similar results. However, a rank-sum based non-parametric test suggested that the signal peptide may not be a strong determinant of the partitioning compared to the transmembrane domain. The observation that the signal peptide is not as strong an indicator for partitioning as the transmembrane domain suggests that signal peptide translocation to the secretory pathway may be a regulated process. Overall, our results indicate that subcellular fractionation and microarray analysis is a powerful tool for characterization of the secretome.

Factors affecting MBP enrichment

While enrichment of transcripts that encode secreted proteins and membrane-bound proteins were evident in the microsomal fraction, some exceptions were noticed. As shown in Figure 3.3A, the distribution of transcripts with a signal peptide suggested

that signal peptide is not the only determinant of partitioning since some transcripts that were predicted to harbor a signal peptide were predominantly in the FP. Some transcripts encoding experimentally verified secreted proteins were also enriched in cytoplasmic fraction compared to microsomal fraction. Similar results were obtained from analysis of the published secretome set for which expression signal was detected from 256 out of 739 secreted proteins (Figure 3.3B). One conspicuous feature was the presence of the hydrophobin gene, MPG1 (MG10315), in the cytosolic fraction from both MM48 and APP samples. Semi-quantitative RT-PCR of samples prepared independently from both the sucrose gradient method and differential centrifugation method confirmed the microarray result (Figure 3.4A and B). RT-PCR was performed on two differently prepared samples. MPG1 has been extensively characterized, and is known to be secreted (Talbot et al., 1996). Hydrophobin proteins play critical roles in fungal development and pathogenicity (Beckerman and Ebbole, 1996; Talbot et al., 1996). The proteins have drawn attention due to their ability to self-assemble and interact with the extracellular matrix (ECM) (Gebbinck et al., 2005). Another example, the extracellular chitinase (MG06594) was also enriched in the FP fraction from MM48. Interestingly, comparison of APP and MM48 microarray data revealed that some transcripts partitioned differently, For example, transcripts of β -glucosidase (MG04582) enriched in the FP fraction in MM48. However, transcripts were enriched in the MBP fraction in APP. The β -glucosidase protein contains a signal peptide according to SignalP analysis. RT-PCR confirmed the microarray results (Figure 3.4C). Thus, partitioning of genes containing a signal peptide might be a regulated process. One possible explanation is the existence of two routes to the secretory pathway as demonstrated in yeast species: co-translational and

post-translational translocation. In higher eukaryotes, most secretory proteins are translocated co-translationally. However, in yeast species, both post-translational and co-translational translocations exist. Possible existence of the two routes has been suggested in filamentous fungi (Conesa et al., 2001). At least in budding yeast, it was noted that the mode of translocation was controlled by the hydrophobicity of signal peptide (Ng et al., 1996).

To test whether fractionation may be affected by the mode of translocation and to see if the hydrophobicity of signal peptide might affect the fractionation of the transcripts, we compared hydropathy plots using the Kyte-Doolittle algorithm (Kyte and Doolittle, 1982). The predicted signal peptide was used as input and the analysis was done using the pepwindow program in EMBOSS package (Rice et al., 2000). The plot suggests that there is little difference between two groups (Figure 3.5). The hydropathy ranged between 1.5 and 2.5 in both groups. Thus, the data suggested that hydrophobicity of the signal peptide may not be a major determinant of routing. At this time, little is known about the secretory pathway in the rice blast fungus. Thus, secretory pathways and the mode of secretion remains to be investigated in future research.

Another possibility to explain differential partitioning is the status of translational activity. If the transcript is not translationally active, the transcript likely will be found in the cytosolic fraction. To investigate the translational activity, polysomal fractionation and Northern analysis were performed. As shown in Figure 3.6, the majority of the BiP transcripts and a putative transcription factor (MG10660) enriched in the FP fraction appeared to be associated with ribosomes. Similarly, MPG1 transcripts were found associated with ribosomes, indicating that MPG1 is translationally active. Western

analysis using epitope-tagged expression strains failed to detect MPG1 protein in mycelia. Taken together, the data suggest that MPG1 transcripts are translationally active, and are likely translocated to the ER post-translationally.

Prediction of GPI-anchored proteins

Among the proteins that are represented outside cell, GPI-anchored proteins are of particular interest. In the rice blast fungus, no GPI-anchored protein has been characterized. Big-Pi analysis yielded 187 predicted GPI-anchored proteins in *M. grisea* (Eisenhaber et al., 2004). GPI-SOM analysis resulted in 305 predicted GPI-anchored proteins (Fankhauser and Maser, 2005b). One hundred thirty one genes were found to be common, and included proteins that function in cell wall metabolism. Analysis of GPI-anchored protein yielded 103 genes with detectible expression in APP. Among these, two encode putative chitin deacetylases (MG01868 and MG08774). Chitin deacetylases play an important role in fungal development, and may contribute to cell wall rigidity (Christodoulidou et al., 1996; Matsuo et al., 2005). Also, in plant pathogenic fungi, modification of the fungal cell wall during infection has been shown (El Gueddari et al., 2002). Thus, it appears that there is differential regulation of chitin metabolism during vegetative growth and infection stage. The role of chitin modification during infection and colonization has not been reported in the rice blast fungus. However, a similar role of chitin modification may operate during infection structure formation in the rice blast fungus. Microarray analyses comparing appressoria and mycelia grown in liquid complete medium and nutrient starvation conditions suggested that expression of these genes was induced during the appressorium formation and carbon starvation (unpublished

data). This prompted us to investigate the alteration of cell wall composition during the infection process. A conidial suspension was incubated on a cellophane film for 24 hrs at room temperature. Using a specific antibody, chitosan was detected in the cell wall of conidia and germ tubes as shown in Figure 3.7. Chitosan was also detected in the narrow infectious hypha. Wheat germ agglutinin (WGA) indicated the presence of chitin in the germ tubes while little staining was noticed in the conidia. Thus, our data suggest that significant alterations of the cell wall occurred both during and after the penetration step. Further analysis of appressoria development on onion epidermal peel revealed stronger staining of the infectious hyphae compared to the hyphae on the surface. RT-PCR analysis indicated that one of chitin deacetylase gene (MG08874) was up-regulated during carbon starvation and late stages of appressoria (Figure 3.7I). We conclude that cell wall alteration occurs during the infectious growth in the rice blast fungus in a manner similar to that observed for other fungi such as *Colletotrichum* and rust fungi (El Gueddari et al., 2002).

Transposable element transcripts are associated with the microsomal fraction

Approximately, 10% of the genome of *M. grisea* is composed of repetitive elements. A large proportion of which are retrotransposons. We were intrigued to observe that genes encoding gag and pol proteins, major constituents of retrotransposons, were abundant in the MBP, particularly in MM48 sample (Figure 3.8). MAGGY, a gypsy-like element is one of the most abundant retrotransposon in the rice blast fungus genome, and active copies exist. It is closely related to the Ty3-element in yeast. Recently, it has been reported that host genes are required for both viral replication and retrotransposon

mobilization, suggesting that retrotransposons utilize the host trafficking machinery in a similar manner to retroviral replication (Irwin et al., 2005). Many stress factors can affect the expression of transposable elements. Transposable elements in yeast such as Ty1, Ty3, and Ty5 are responsive to pheromone and nutrients. Ty3 LTR encodes the binding site for Msn2 and Msn4, transcription factors that respond to stress and carbon starvation. The analysis of MAGGY LTR with TESS software revealed a potential binding site for nit-2 (URL: <http://www.cbil.upenn.edu/tess>). No comprehensive study regarding the expression of MAGGY in *M. grisea* has been reported. However, a recent study suggested that a variety of stresses such as heat shock and oxidative stress may affect the transcription of MAGGY (Ikeda et al., 2001). Thus, it appears that possible similar mechanisms of mobilization and trafficking operate for the propagation of transposable elements in filamentous fungi and yeast.

Discussion

In this article, we performed a detailed analysis of mRNA following microsomal fractionation for the rice blast fungus using DNA microarrays. Genes that encode proteins which are secreted and/or localize to the membrane play crucial roles during pathogenesis such as surface sensing, degradation of host cell walls, and nutrient uptake. To define these proteins, we conducted microarray experiments using microsome-associated RNA and free cytosolic RNA. The enrichment of transcripts in each fraction derived from different growth conditions suggests that the enrichment may be affected by many endogenous factors such as transcription level, translation level, post-translational modification, as well as physical structure of proteins. Our results show that proteins with

a signal peptide appear to be routed to the secretory pathway by two different modes (co-translational and post-translational translocation) in the rice blast fungus. The existence of two translocation mechanisms has been previously postulated to exist in filamentous fungi (Conesa et al., 2001). Diehn et al noted that many types of protein modification are coupled to translation that may affect localization (Diehn et al., 2000). These modifications include proteolytic cleavage by signal peptidases, myristoylation of N-terminus of proteins. In higher eukaryotes, secretory proteins are translocated while they are being translated (co-translational translocation). However, there are two modes of translocation in lower eukaryotes including budding yeast. If a protein takes a route to post-translational translocation, and/or if regulation is at the level of initiation of translation, most of the transcripts will be detected in the cytosolic fraction. In our study, we found no relationship between signal peptide and hydrophobicity for predicting the mechanism of entering the secretory pathway. In *Yarrowia lipolytica*, the routing mechanism of alkaline extracellular protease can be switched without altering the overall hydrophobicity of the signal peptide (Matoba and Ogrydziak, 1998). The authors suggested that the factors other than the hydrophobicity of signal peptide might contribute to the routing such as conformation and orientation of the signal peptide. Another possible explanation of enrichment of transcripts contrary to our expectation might be due to incomplete annotations of the genes. Gene annotation is an ongoing process, necessitating experimental verification. In fact, Toyoda et al showed that, using a similar approach, incomplete sequence information from EST data can be corrected, and they were able to identify a novel seven transmembrane receptor in humans (Toyoda et al., 2003). Also, transcripts might be associated with subcellular components (Diehn et al.,

2000; Shepard et al., 2003). Fractionation can be affected by other targeting sequences such as mitochondrial targeting sequences and its length (Marc et al., 2002; Sylvestre et al., 2003). A translational profiling experiment in budding yeast suggested that there is an inverse relationship between transcript length and ribosomal association (Arava et al., 2003). As in Figure 6, MPG1 transcripts appeared to be fractionated in a higher polysome fraction compared to the BiP, consistent with this observation. However, overall for our data, we did not find a strong correlation between sequence length and targeting (data not shown).

In our work, we observed numerous genes that exhibited different localization patterns between MM48 and APP conditions. Transcripts for some genes were enriched in one fraction under one growth condition, but were localized to the other fraction in the other growth condition. One of the examples we were able to confirm was β -glucosidase. β -glucosidase has a predicted signal peptide. In the APP condition, the transcript localized to the microsomal fraction while the transcript was mainly found in cytosolic fraction in MM48. This supports the hypothesis that signal peptide mediated secretion is a regulated process.

Chitin is an important component of the fungal cell wall. Also, chitin oligomers and chitosan are strong inducers of plant responses. It has been suggested that chitinous material is converted into less detectible material by chitinase enzymes from the host during ectomycorrhizal colonization in order to escape activation of host defenses. Thus, the hemibiotrophic fungus, *M. grisea* may have developed a similar strategy to avoid host responses during penetration. The presence of a putative GPI-anchor in a chitin deacetylase appears to be an efficient strategy to confine the enzyme to a specific

location without being lost to the surrounding environment. Localization data of the chitin deacetylases from diverse fungi suggest the presence of the proteins in periplasm and extracellular environment (Tsigos et al., 2000), consistent with the presence of a GPI-anchor. Microarray data suggest that chitin deacetylases are up-regulated by the deprivation of carbon nutrient source and in appressoria compared to mycelia grown in liquid complete medium as shown in Figure 7. In previous work, a putative chitin deacetylase was found to interact with adenylate cyclase (MAC1) in the rice blast fungus (Kulkarni and Dean, 2004), suggesting that its activity might be regulated by the cAMP pathway. cAMP is an inducer of appressorium formation in the rice blast fungus, and is a central messenger in sensing nutrient starvation. Thus, chitin deacetylases will be an interesting subject for functional analysis in future research.

In our analysis, the microsomal fraction was rich in components from transposable elements. Early studies suggested that retroviruses are associated with the microsomal fraction (Gielkens et al., 1974; Vecchio et al., 1973). Compared to research on yeast transposable elements, little is known regarding the expression and regulation of transposable elements in filamentous fungi. The replication of Ty3 in yeast requires intracellular trafficking (Irwin et al., 2005). Hypovirulence associated virus (HAV) in *Cryphonectria parasitica* is also known to require membranes for replication (Fahima et al., 1993; Hansen et al., 1985). In addition, recent studies show that transposable elements are transcriptionally regulated in response to external stresses (Ikeda et al., 2001). In yeast, Ty elements are responsive to various signals including pheromone and nutrients. In our work, we observed that LTR may serve as a promoter since it contains a predicted nit-2 binding site. NUT1, the *M. grisea* homolog of nit-2, is a global nitrogen

regulator (Froeliger and Carpenter, 1996). Thus, MAGGY might be transcriptionally regulated by the availability of nitrogen. Other data suggest that transposable elements in the rice blast fungus are regulated by nutrient availability and developmental stage.

In this article, we employed microarray analysis to define the secreted and membrane-bound proteins in the rice blast fungus. Our results showed that the microarray data is robust, and led to a more detailed evaluation of transcript localization. Our data also suggests that many physiological factors, one being translational, regulate localization. Thus, interpretation of localization studies should be addressed with some caution. Our work revealed that a simple comparison of microsome-associated and cytosolic RNA by microarray analysis to identify secreted proteins is insufficient. We found that factors such as routing path to ER and translational control might lead to unanticipated observations. However, the data do provide evidence for the fine regulation of secretion in fungi. Thus, such data can be used to help define the routing mechanisms that determine the secretory pathway. Also, the use of the approach explored in this research for a variety of growth and developmental conditions will provide a compendium of potential secreted and membrane-bound proteins. Such information will greatly aid functional annotation of genes with unknown function.

References

1. Arava, Y., Wang, Y., Storey, J.D., Liu, C.L., Brown, P.O., and Herschlag, D. (2003). Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci U. S. A* *100*, 3889-3894.
2. Beckerman, J.L. and Ebbole, D.J. (1996). MPG1, a gene encoding a fungal hydrophobin of *Magnaporthe grisea*, is involved in surface recognition. *Mol Plant Microbe Interact.* *9*, 450-456.
3. Bendtsen, J.D., Nielsen, H., von Heijne, G., and Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* *340*, 783-795.
4. Chiou, C.H., Lee, L.W., Owens, S.A., Whallon, J.H., Klomparens, K.L., Townsend, C.A., and Linz, J.E. (2004). Distribution and sub-cellular localization of the aflatoxin enzyme versicolorin B synthase in time-fractionated colonies of *Aspergillus parasiticus*. *Archives of Microbiology* *182*, 67-79.
5. Christodoulidou, A., Bouriotis, V., and Thireos, G. (1996). Two sporulation-specific chitin deacetylase-encoding genes are required for the ascospore wall rigidity of *Saccharomyces cerevisiae*. *J Biol Chem.* *271*, 31420-31425.
6. Clergeot, P.H., Gourgues, M., Cots, J., Laurans, F., Latorse, M.P., Pepin, R., Tharreau, D., Notteghem, J.L., and Lebrun, M.H. (2001). PLS1, a gene encoding a

tetraspanin-like protein, is required for penetration of rice leaf by the fungal pathogen *Magnaporthe grisea*. Proc. Natl Acad. Sci U. S. A 98, 6963-6968.

7. Conesa,A., Punt,P.J., van Luijk,N., and van den Hondel,C.A. (2001). The secretion pathway in filamentous fungi: a biotechnological view. Fungal Genet Biol 33, 155-171.
8. Dean,R.A. (1997). Signal pathways and appressorium morphogenesis. Annu. Rev. Phytopathol. 35, 211-234.
9. Dean,R.A., Talbot,N.J., Ebbole,D.J., Farman,M.L., Mitchell,T.K., Orbach,M.J., Thon,M., Kulkarni,R., Xu,J.R., Pan,H., Read,N.D., Lee,Y.H., Carbone,I., Brown,D., Oh,Y.Y., Donofrio,N., Jeong,J.S., Soanes,D.M., Djonovic,S., Kolomiets,E., Rehmeyer,C., Li,W., Harding,M., Kim,S., Lebrun,M.H., Bohnert,H., Coughlan,S., Butler,J., Calvo,S., Ma,L.J., Nicol,R., Purcell,S., Nusbaum,C., Galagan,J.E., and Birren,B.W. (2005). The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature 434, 980-986.
10. DeZwaan,T.M., Carroll,A.M., Valent,B., and Sweigard,J.A. (1999). *Magnaporthe grisea* Pth11p is a novel plasma membrane protein that mediates appressorium differentiation in response to inductive surface cues. Plant Cell 11, 2012-2030.

11. Diehn,M., Eisen,M.B., Botstein,D., and Brown,P.O. (2000). Large-scale identification of secreted and membrane-associated gene products using DNA microarrays. *Nat Genet* 25, 58-62.
12. Eisenhaber,B., Schneider,G., Wildpaner,M., and Eisenhaber,F. (2004). A sensitive predictor for potential GPI lipid modification sites in fungal protein sequences and its application to genome-wide studies for *Aspergillus nidulans*, *Candida albicans*, *Neurospora crassa*, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *J Mol Biol* 337, 243-253.
13. El Gueddari,N.E., Rauchhaus,U., Moerschbacher,B.M., and Deising,H.B. (2002). Developmentally regulated conversion of surface-exposed chitin to chitosan in cell walls of plant pathogenic fungi. *New Phytologist* 156, 103-112.
14. Fahima,T., Kazmierczak,P., Hansen,D.R., Pfeiffer,P., and Van Alfen,N.K. (1993). Membrane-associated replication of an unencapsidated double-strand RNA of the fungus, *Cryphonectria parasitica* . *Virology* 195, 81-89.
15. Fankhauser,N. and Maser,P. (2005a). Identification of GPI anchor attachment signals by a Kohonen self-organizing map. *Bioinformatics* 21, 1846-1852.
16. Fankhauser,N. and Maser,P. (2005b). Identification of GPI anchor attachment signals by a Kohonen self-organizing map. *Bioinformatics* 21, 1846-1852.

17. Foster,A.J., Jenkinson,J.M., and Talbot,N.J. (2003). Trehalose synthesis and metabolism are required at different stages of plant infection by *Magnaporthe grisea*. EMBO J 22, 225-235.
18. Froeliger,E.H. and Carpenter,B.E. (1996). NUT1, a major nitrogen regulatory gene in *Magnaporthe grisea*, is dispensable for pathogenicity. Mol Gen Genet. 251, 647-656.
19. Fudal,I., Bohnert,H.U., Tharreau,D., and Lebrun,M.H. (2005). Transposition of MINE, a composite retrotransposon, in the avirulence gene ACE1 of the rice blast fungus *Magnaporthe grisea*. Fungal Genet. Biol.
20. Galagan,J.E., Calvo,S.E., Borkovich,K.A., Selker,E.U., Read,N.D., Jaffe,D., FitzHugh,W., Ma,L.J., Smirnov,S., Purcell,S., Rehman,B., Elkins,T., Engels,R., Wang,S., Nielsen,C.B., Butler,J., Endrizzi,M., Qui,D., Ianakiev,P., Bell-Pedersen,D., Nelson,M.A., Werner-Washburne,M., Selitrennikoff,C.P., Kinsey,J.A., Braun,E.L., Zelter,A., Schulte,U., Kothe,G.O., Jedd,G., Mewes,W., Staben,C., Marcotte,E., Greenberg,D., Roy,A., Foley,K., Naylor,J., Stange-Thomann,N., Barrett,R., Gnerre,S., Kamal,M., Kamvysselis,M., Mauceli,E., Bielke,C., Rudd,S., Frishman,D., Krystofova,S., Rasmussen,C., Metzenberg,R.L., Perkins,D.D., Kroken,S., Cogoni,C., Macino,G., Catcheside,D., Li,W., Pratt,R.J., Osmani,S.A., DeSouza,C.P.C., Glass,L., Orbach,M.J., Berglund,J.A., Voelker,R., Yarden,O., Plamann,M., Seiler,S., Dunlap,J., Radford,A., Aramayo,R., Natvig,D.O., Alex,L.A., Mannhaupt,G., Ebbole,D.J., Freitag,M., Paulsen,I., Sachs,M.S., Lander,E.S.,

- Nusbaum,C., and Birren,B. (2003). The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422, 859-868.
21. Gebbink,M.F., Claessen,D., Bouma,B., Dijkhuizen,L., and Wosten,H.A. (2005). Amyloids--a functional coat for microorganisms. *Nat Rev Microbiol* 3, 333-341.
 22. Gielkens,A.L., Salden,M.H., and Bloemendal,H. (1974). Virus-specific messenger RNA on free and membrane-bound polyribosomes from cells infected with Rauscher leukemia virus. *Proc. Natl Acad. Sci U. S. A* 71, 1093-1097.
 23. Gilbert,R.D., Johnson,A.M., and Dean,R.A. (1996). Chemical signals responsible for appressorium formation in the rice blast fungus *Magnaporthe grisea*. *Physiol Mol Plant P* 48, 335-346.
 24. Gygi,S.P., Rochon,Y., Franza,B.R., and Aebersold,R. (1999). Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* 19, 1720-1730.
 25. Hamer,J.E., Howard,R.J., Chumley,F.G., and Valent,B. (1988). A mechanism for surface attachment in spores of a plant pathogenic fungus. *Science* 239, 288-290.
 26. Hansen,D.R., Van Alfen,N.K., GILLIES,K., and POWELL,W.A. (1985). Naked dsRNA associated with hypovirulence of *Endothia parasitica* is packaged in fungal vesicles. *J Gen Virol* 66, 2605-2614.

27. Howard,R.J., Ferrari,M.A., Roach,D.H., and Money,N.P. (1991). Penetration of hard substrates by a fungus employing enormous turgor pressures. Proc. Natl Acad. Sci U. S. A 88, 11281-11284.
28. Ikeda,K., Nakayashiki,H., Takagi,M., Tosa,Y., and Mayama,S. (2001). Heat shock, copper sulfate and oxidative stress activate the retrotransposon *MAGGY* resident in the plant pathogenic fungus *Magnaporthe grisea*. Mol Genet Genomics 266, 318-325.
29. Irwin,B., Aye,M., Baldi,P., Beliakova-Bethell,N., Cheng,H., Dou,Y., Liou,W., and Sandmeyer,S. (2005). Retroviruses and yeast retrotransposons use overlapping sets of host genes. Genome Res. 15, 641-654.
30. Kamakura,T., Yamaguchi,S., Saitoh,K., Teraoka,T., and Yamaguchi,I. (2002). A novel gene, CBP1, encoding a putative extracellular chitin-binding protein, may play an important role in the hydrophobic surface sensing of *Magnaporthe grisea* during appressorium differentiation. Mol Plant Microbe Interact. 15, 437-444.
31. Kang,S., Lebrun,M.H., Farrall,L., and Valent,B. (2001). Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. Mol Plant Microbe Interact. 14, 671-674.

32. Kihara,J., Sato,A., Okajima,S., and Kumagai,T. (2001). Molecular cloning, sequence analysis and expression of a novel gene induced by near-UV light in *Bipolaris oryzae*. *Mol Genet Genomics* 266, 64-71.
33. Koczynski,C.C., Noordermeer,J.N., Serano,T.L., Chen,W.Y., Pendleton,J.D., Lewis,S., Goodman,C.S., and Rubin,G.M. (1998). A high throughput screen to identify secreted and transmembrane proteins involved in *Drosophila* embryogenesis. *PNAS* 95, 9973-9978.
34. Krogh,A., Larsson,B., von Heijne,G., and Sonnhammer,E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305, 567-580.
35. Kuhn,K.M., DeRisi,J.L., Brown,P.O., and Sarnow,P. (2001). Global and specific translational regulation in the genomic response of *Saccharomyces cerevisiae* to a rapid transfer from a fermentable to a nonfermentable carbon source. *Mol. Cell. Biol.* 21, 916-927.
36. Kulkarni,R.D. and Dean,R.A. (2004). Identification of proteins that interact with two regulators of appressorium development, adenylate cyclase and cAMP-dependent protein kinase A, in the rice blast fungus *Magnaporthe grisea*. *Mol Genet. Genomics* 270, 497-508.

37. Kyte, J. and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157, 105-132.
38. Lauge, R. and de Wit, P.J. (1998). Fungal avirulence genes: structure and possible functions. *Fungal Genet Biol* 24, 285-297.
39. Marc, P., Margeot, A., Devaux, F., Blugeon, C., Corral-Debrinski, M., and Jacq, C. (2002). Genome-wide analysis of mRNAs targeted to yeast mitochondria. *EMBO Rep.* 3, 159-164.
40. Matoba, S. and Ogrydziak, D.M. (1998). Another factor besides hydrophobicity can affect signal peptide interaction with signal recognition particle. *J Biol Chem.* 273, 18841-18847.
41. Matsuo, Y., Tanaka, K., Matsuda, H., and Kawamukai, M. (2005). *cda1+*, encoding chitin deacetylase is required for proper spore formation in *Schizosaccharomyces pombe*. *FEBS Lett.* 579, 2737-2743.
42. Mechler, B.M. (1987). Isolation of messenger RNA from membrane-bound polysomes. *Methods Enzymol.* 152, 241-248.
43. Mikulits, W., Pradet-Balade, B., Habermann, B., Beug, H., Garcia-Sanz, J.A., and Mullner, E.W. (2000). Isolation of translationally controlled mRNAs by differential screening. *FASEB J* 14, 1641-1652.

44. Ng,D.T., Brown,J.D., and Walter,P. (1996). Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J Cell Biol* *134*, 269-278.
45. Pradet-Balade,B., Boulme,F., Beug,H., Mullner,E.W., and Garcia-Sanz,J.A. (2001). Translation control: bridging the gap between genomics and proteomics? *Trends Biochem Sci* *26*, 225-229.
46. Rice,P., Longden,I., and Bleasby,A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* *16*, 276-277.
47. Schafer,W. (1994). Molecular mechanisms of fungal pathogenicity to plants. *Annu. Rev Phytopathol.* *32*, 461-477.
48. Sesma,A. and Osbourn,A.E. (2004). The rice blast pathogen undergoes developmental processes typical of root-infecting fungi. *Nature* *431*, 582-586.
49. Shepard,K.A., Gerber,A.P., Jambhekar,A., Takizawa,P.A., Brown,P.O., Herschlag,D., DeRisi,J.L., and Vale,R.D. (2003). Widespread cytoplasmic mRNA transport in yeast: Identification of 22 bud-localized transcripts using DNA microarray analysis. *Proc. Natl. Acad. Sci U. S. A* *100*, 11429-11434.
50. Stoltenburg,R., Wartmann,T., Kunze,I., and Kunze,G. (1995). Reliable method to prepare RNA from free and membrane-bound polysomes from different yeast species. *Biotechniques* *18*, 564-6, 568.

51. Sweigard,J.A., Chumley,F.G., and Valent,B. (1992). Disruption of a *Magnaporthe grisea* cutinase gene. *Molecular Genetics and Genomics (Historical Archive)* 232, 183-190.
52. Sylvestre,J., Vialette,S., Corral,D.M., and Jacq,C. (2003). Long mRNAs coding for yeast mitochondrial proteins of prokaryotic origin preferentially localize to the vicinity of mitochondria. *Genome Biol* 4, R44.
53. Talbot,N.J., Kershaw,M.J., Wakley,G.E., De Vries,O., Wessels,J., and Hamer,J.E. (1996). MPG1 encodes a fungal hydrophobin involved in surface interactions during infection-related development of *Magnaporthe grisea*. *Plant Cell* 8, 985-999.
54. Toyoda,N., Nagai,S., Terashima,Y., Motomura,K., Haino,M., Hashimoto,S.i., Takizawa,H., and Matsushima,K. (2003). Analysis of mRNA with microsomal fractionation using a SAGE-Based DNA microarray system facilitates identification of the genes encoding secretory proteins. *Genome Res.* 13, 1728-1736.
55. Tsigos,I., Martinou,A., Kafetzopoulos,D., and Bouriotis,V. (2000). Chitin deacetylases: new, versatile tools in biotechnology. *Trends Biotechnol.* 18, 305-312.
56. Vecchio,G., Tsuchida,N., Shanmugam,G., and Green,M. (1973). Virus-specific messenger RNA and nascent polypeptides in polyribosomes of cells replicating murine sarcoma-leukemia viruses. *Proc. Natl Acad. Sci U. S. A* 70, 2064-2068.

57. Walton,J.D. (1994). Deconstructing the cell wall. *Plant Physiol* *104*, 1113-1118.

58. Wu,S.C., Ham,K.S., Darvill,A.G., and Albersheim,P. (1997). Deletion of two endo-beta-1,4-xylanase genes reveals additional isozymes secreted by the rice blast fungus. *Mol Plant Microbe Interact.* *10*, 700-708.

Table 3.1 Primers used for RT-PCR analysis

Primers	Sequence
MG01472F	5'-ATGGCTGCGCTCGCAAGCAA-3'
MG01472R	5'-GCTGCCGAGATAGCTGCGAA-3'
MG01931F	5'-ATGACCTCCAACGCAGCTCC-3'
MG01931R	5'-TCCTCATCAACCTTCCTGACG-3'
MG02503F	5'-CCTCGCGTTTGTCAAGACCG-3'
MG02503R	5'-TAGCTCATCGTGGGAACGAGG-3'
MG02647F	5'-ATGCAGTTCTCCATCACCAA-3'
MG02647R	5'-GACACGGAGGGCGGCGCG-3'
MG03416F	5'-ATGACTGCCGCTAGAGGTCC-3'
MG03416R	5'-TCGCAAATATGTGCTATCCACCC-3'
MG05110F	5'-TTGTACAAAAAAGCAGGCTTAATGAGATGGCTTTTGGTGGAC-3'
MG05110R	5'-TTTGTACAAGAAAGCTGGGTCAATTGTCTTGGCCTTGTGCTT-3'
MG06432F	5'-GTTACGAAGAACAGGCGAACC-3'
MG06432R	5'-TTATGGGGGTATCGACGGTGG-3'
MG06594F	5'-ATGAAGTTTTCACTCTCTGCTTTGG-3'
MG06594R	5'-AGCCTTGCACTTCTTCTTCAGG-3'
MG07975F	5'-ATGCAGATCCCCACCGTCTG-3'
MG07975R	5'-CGCAAAGAAGCTGCCACC-3'
MG08164F	5'-ATGGGTAACCTTAGGAGCCT-3'
MG08164R	5'-CAACTCGTCCTTGCCAGTGA-3'
MG08264F	5'-TTGTACAAAAAAGCAGGCTTAATGCAGCTCGCAACAGTTTTTC-3'
MG08264R	5'-TTTGTACAAGAAAGCTGGGTCAATTTTGGTCTTGGAATCTCCC-3'
MG09736F	5'-ATGGTTTCCTCTCTGGCCATC-3'
MG09736R	5'-CAAAGCCAGGACAGCGAGAG-3'
MG10315F	5'-GAAGAACCCGGGGATCCCCGCCCTGGTGAG-3'
MG10315R	5'-GAAGAAGTGCAGGTCAAGGGGATAGGAGTGC-3'
MG10660F	5'-ATGGCGCGGGTGAACACATTC-3'
MG10660R	5'-CTTCTTCCCATTAGCATCTCC-3'

Table 3.2 List of Top 25 transcripts enriched in MBP and FP fraction in MM48 condition

(A) MBP enriched transcripts

ID ^a	MBP/FP ^b	MBP ^c	Description
AZM04039	12.967709	221.125	(MG00565.1) predicted protein
AZM13403	5.55403	14397.7	heat shock 70kD protein 5 [<i>Rattus norvegicus</i>] (BIP)
AZM05215	4.633237	926.7205	NADPH-cytochrome P450 reductase (CPR) [<i>A. niger</i>]
AZM12287	4.50943	5211.9927	delta-9 fatty acid desaturase[<i>Ajellomyces capsulatus</i>]
AZM08774	4.1050754	2897.955	predicted protein [<i>N. crassa</i>]
AZM11244	4.0950184	247.23325	hypothetical protein [<i>N. crassa</i>]
AZM08292	4.090344	416.193	predicted protein [<i>N. crassa</i>]
AZM06431	4.074249	746.806	(MG05689.1) hypothetical protein
AZM08528	3.9035997	254.563	Dcr2p [<i>S. cerevisiae</i>]
AZM04055	3.882538	268.92877	DNA J protein [<i>S. pombe</i>]
AZM02769	3.855157	349.138	putative calcium P-type ATPase [<i>N. crassa</i>]
AZM03813	3.8466437	558.72754	hypothetical protein [<i>N. crassa</i>]
AZM00220	3.71775	695.4177	isoamyl alcohol oxidase [<i>A. oryzae</i>]
AZM07293	3.716761	384.702	(MG05746.1) hypothetical protein
AZM06777	3.6144938	2596.335	CAN1 [<i>S. castellii</i>]
AZM04911	3.5928197	100.09475	probable zinc metallo-protease [<i>N. crassa</i>]
AZM04843	3.5545497	2095.9175	yeast ero1 homolog [<i>S. pombe</i>]
AZM00221	3.543917	297.431	putative 5' nucleotidase family protein [<i>S. pombe</i>]
AZM04628	3.4986262	6418.097	related to ubiquitin-activating enzyme UBA2 [<i>N. crassa</i>]
AZM08512	3.4897518	1475.185	conserved hypothetical protein [<i>N. crassa</i>] related to n-alkane-inducible cytochrome P450 [<i>N. crassa</i>]
AZM00174	3.4377732	30104.648	
AZM12598	3.4184365	355.2115	oleate delta-12 desaturase [<i>E. nidulans</i>]
AZM11813	3.387078	158.88324	membrane drug transporter [<i>S. pombe</i>]
AZM12761	3.3712056	335.04047	hypothetical protein [<i>N. crassa</i>]
AZM03424	3.3671718	567.03174	predicted protein [<i>N. crassa</i>]

Table 3.2 (continued)**(B) FP-enriched transcripts**

ID ^a	MBP/FP ^b	MBP ^c	Description
AZM07835	0.17259425	277.4985	(MG08817.1) predicted protein
AZM08794	0.19288206	395.5735	(MG09385.1) predicted protein
AZM10704	0.1981571	2749.0352	(MG10660.1) predicted transcription factor
AZM06557	0.19855645	5802.6924	(MG07975.1) predicted protein
AZM07319	0.217739	753.1502	(MG08480.1) hypothetical protein
AZM01795	0.21867003	220.53125	(MG03865.1) predicted protein
AZM04220	0.22324422	551.4535	(MG05405.1) predicted protein
AZM08632	0.22492199	651.1362	(MG09270.1) predicted protein
AZM09190	0.2270839	447.98776	predicted protein [<i>N. crassa</i>]
AZM06551	0.23422275	314.9025	(MG07969.1) predicted protein
AZM07407	0.2370141	127.2575	(MG08532.1) predicted protein
AZM11324	0.2432253	1782.26	(MG11035.1) predicted protein
AZM06044	0.25260064	287.37402	(MG07632.1) hypothetical protein
AZM12877	0.2526331	897.9897	(MG02165.1) predicted protein
AZM10225	0.2596041	401.677	MPG1 precursor hydrophobin-like protein [<i>M. grisea</i>]
AZM06674	0.26697424	2006.2825	predicted protein [<i>N. crassa</i>]
AZM08603	0.26746973	1133.4222	predicted protein
AZM11078	0.27011558	8100.155	(MG01197.1) predicted protein
AZM00435	0.27707633	532.6305	predicted protein
AZM05701	0.27752396	3592.4749	(MG05638.1) predicted protein
AZM04545	0.28288916	204.08275	(MG05556.1) predicted protein
AZM09326	0.2866588	11511.732	(MG09736.1) predicted protein
AZM08800	0.28800836	417.5095	(MG09389.1) predicted protein
AZM05302	0.28830865	236.198	(MG07177.1) predicted protein
AZM13533	0.2901562	209.91699	predicted protein

^aAgilent probe ID^bRatio of MBP/FP^cMBP signal intensity

Table 3.3 List of Top 25 transcripts enriched in MBP and FP fraction in APP condition

(A) MBP-enriched transcripts

ID ^a	MBP/FP ^b	MBP ^c	Description
AZM07039	19.78929	18560.6	(MG08296.1) predicted protein
AZM13403	6.227844	5289.79	heat shock 70kD protein 5 [<i>Rattus norvegicus</i>] (BIP)
AZM03638	6.054921	23819.85	versicolorin B synthase [<i>A. parasiticus</i>]
AZM01518	5.7955	23867.75	Subtilisin-like proteinase Spm1 (Serine protease of M. 1)
AZM11244	5.761551	229.0108	hypothetical protein [<i>N. crassa</i>]
AZM03716	5.290731	2958.905	(MG05110.1) predicted protein
AZM06278	5.177973	6399.167	(MG07790.1) hypothetical protein
AZM07391	5.005079	6298.96	Protein disulfide isomerase precursor [<i>Humicola insolens</i>]
AZM06094	4.721647	1706.587	serine-type carboxypeptidase homolog [<i>A. nidulans</i>]
AZM03451	4.719541	198.7645	(MG04925.1) predicted protein
AZM03373	4.67437	4596.95	putative membrane glycoprotein; possible vacuolar protein sorting /targeting [<i>S. pombe</i>]
AZM06847	4.644655	1471.14	probable protein disulfide-isomerase precursor [<i>N. crassa</i>]
AZM06207	4.448503	1176.006	dipeptidyl aminopeptidase type IV [<i>A. niger</i>]
AZM11014	4.408891	3350.708	(MG10869.1) predicted protein
AZM00221	4.406614	423.8058	putative 5' nucleotidase family protein [<i>S. pombe</i>]
AZM13388	4.325295	1895.057	cephalosporin esterase [<i>Rhodosporidium toruloides</i>]
AZM10002	4.310957	1676.503	(MG10168.1) predicted protein
AZM08774	4.296306	718.8707	predicted protein [<i>N. crassa</i>]
AZM09260	4.256019	717.5427	Aminopeptidase Y precursor [<i>S. cerevisiae</i>]
AZM07633	4.188554	1600.765	Glycolipid 2-alpha-mannosyltransferase [<i>Candida albicans</i>]
AZM01156	4.143649	337.2945	hypothetical protein [<i>N. crassa</i>]
AZM12034	4.104717	7817.407	calnexin [<i>A. niger</i>]
AZM08292	4.081607	344.7683	predicted protein [<i>N. crassa</i>]
AZM09149	4.064821	13695.54	(MG09618.1) predicted protein
AZM05267	4.050938	885.689	(MG07153.1) predicted protein

Table 3.3 (continued)**(B) FP-enriched transcripts**

ID ^a	MBP/FP ^b	MBP ^c	Description
AZM06992	0.100331	1384.5995	(MG08264.1) predicted protein
AZM10225	0.1005621	568.0483	MPG1 precursor hydrophobin-like protein [<i>M. grisea</i>]
AZM04518	0.1246823	107.69645	(MG06665.1) predicted protein
AZM07319	0.1351433	586.162	(MG08480.1) hypothetical protein
AZM12924	0.1781336	156.41525	secreted aspartic proteinase precursor
AZM09729	0.2163313	505.1833	(MG10010.1) predicted protein
AZM12877	0.2279341	492.04922	(MG02165.1) predicted protein
AZM06557	0.2304961	1221.6002	(MG07975.1) predicted protein
AZM02176	0.2365489	171.53075	YOL080c [<i>S. cerevisiae</i>]
AZM10796	0.2365564	336.37726	(MG10705.1) predicted protein
AZM07407	0.240192	132.29425	(MG08532.1) predicted protein
AZM07827	0.245085	571.3735	oxidoreductase, aldo/keto reductase[<i>E. faecalis</i> V583]
AZM02359	0.2483728	315.29724	reverse transcriptase [<i>M. grisea</i>]
AZM06674	0.249723	240.58202	predicted protein [<i>N. crassa</i>]
AZM06521	0.250198	288.88998	(MG07951.1) predicted protein
AZM03121	0.2610821	1583.2899	endochitinase [<i>Amanita muscaria</i>]
AZM12465	0.261983	306.15497	predicted protein [<i>N. crassa</i>]
AZM11078	0.2634089	7221.6597	(MG01197.1) predicted protein
AZM09190	0.2639213	118.02548	predicted protein [<i>N. crassa</i>]
AZM09130	0.266747	1276.148	(MG09610.1) hypothetical protein
AZM11324	0.2739945	681.0815	(MG11035.1) predicted protein
AZM07369	0.2836464	267.172	(MG08506.1) predicted protein
AZM12252	0.2927858	432.838	clathrin coat assembly protein [<i>S. pombe</i>]
AZM12638	0.3039556	731.7153	predicted protein [<i>N. crassa</i>]
AZM10704	0.3187566	7121.4653	(MG10660.1) predicted transcription factor

^aAgilent probe ID^bRatio of MBP/FP^cMBP signal intensity

Table 3.4 Statistical significance for signal ratio for signal peptide and transmembrane domains

A. Mean signal ratio

Samples	SIG+ ^a	SIG- ^b	TM+ ^c	TM- ^d
MM48	1.12823	1.07479	1.31374	1.02616
APP	1.39492	1.08706	1.51451	1.04043

B. Statistical analysis of signal ratio on signal peptide and TM domains

Samples	p value ^e	Non-parametric p ^f
MM48 (SP)	0.0004	0.3371
MM48 (TM)	0	<0.0001
APP (SP)	0	<0.0001
APP (TM)	0	<0.0001

^{a,b} Mean signal ratios (MBP/FP) of transcripts with (SIG+) or without (SIG-) signal peptide, respectively.

^{c,d} Mean signal ratios (MBP/FP) of transcripts with (TM+) or without (TM-) transmembrane domains, respectively.

^e p-values of t-test (assuming unequal variance) when transcripts are grouped by the presence or absence of a signal peptide.

^f p-values of rank-sum-based non-parametric test when transcripts are grouped by the presence or absence of a transmembrane domains.

Table 3.5 List of GPI-anchored proteins in the rice blast fungus genome

ID ^a	MGID ^b	Description
AZM09074	MG09575	hypothetical protein [<i>N. crassa</i>]
AZM09067	MG09570	(MG09570.1) predicted protein
AZM09326	MG09736	(MG09736.1) predicted protein
AZM09322	MG01051	related to PET8 protein [<i>N. crassa</i>]
AZM09189	MG09640	Taka-amylase A (Taa-G1) precursor
AZM08590	MG01009	(MG01009.1) predicted protein
AZM08913	MG09460	(MG09460.1) predicted protein
AZM08807	MG09395	(MG09395.1) predicted protein
AZM08844	MG09412	predicted protein [<i>N. crassa</i>]
AZM09982	MG05913	(MG05913.1) predicted protein
AZM10171	MG01134	UTR2 PROTEIN (UNKNOWN TRANSCRIPT 2 PROTEIN)
AZM10069	MG05483	putative membrane-associated or secreted protein [<i>M. grisea</i>]
AZM10050	MG10206	(MG10206.1) predicted protein
AZM09574	MG09918	rAsp f 9 [<i>A. fumigatus</i>]
AZM09431	MG09804	(MG09804.1) predicted protein
AZM09476	MG09840	hypothetical protein [<i>N. crassa</i>]
AZM07639	MG08698	Candidapepsin 8 precursor (Aspartate protease 8) [<i>C. albicans</i>]
AZM07152	MG08370	beta (1-3) glucanosyltransferase Gel3p [<i>A. fumigatus</i>]
AZM07129	MG08354	(MG08354.1) predicted protein
AZM07126	MG08351	(MG08351.1) predicted protein
AZM08289	MG09079	hypothetical protein B24B19.340 [<i>N. crassa</i>]
AZM08206	MG09054	(MG09054.1) predicted protein
AZM08364	MG09115	(MG09115.1) predicted protein
AZM07829	MG08812	(MG08812.1) predicted protein
AZM07828	MG08811	(MG08811.1) predicted protein
AZM07843	MG08824	(MG08824.1) predicted protein
AZM07841	MG08823	predicted protein [<i>N. crassa</i>]
AZM07814	MG08798	(MG08798.1) predicted protein
AZM07711	MG08754	(MG08754.1) predicted protein
AZM07781	MG08774	chitin deacetylase [<i>Blumeria graminis</i>]
AZM07775	MG05464	endo-beta-1,4-D-xylanase [<i>M. grisea</i>]
AZM12395	MG01861	(MG01861.1) predicted protein
AZM12410	MG01872	(MG01872.1) predicted protein
AZM12405	MG01868	chitin deacetylase [<i>Blumeria graminis</i>]
AZM12795	MG02112	(MG02112.1) predicted protein
AZM12787	MG02106	(MG02106.1) predicted protein
AZM12134	MG01678	infection structure specific protein [<i>M. grisea</i>]
AZM12036	MG01609	(MG01609.1) predicted protein
AZM12030	MG01603	predicted protein [<i>N. crassa</i>]
AZM12231	MG01750	(MG01750.1) predicted protein
AZM13388	MG02497	cephalosporin esterase precursor [<i>Rhodospiridium toruloides</i>]
AZM13295	MG02431	(MG02431.1) predicted protein
AZM13602	MG06189	probable cell wall biogenesis protein - fission yeast [<i>S. pombe</i>]

Table 3.5 (continued)

AZM13184	MG00210	predicted protein [<i>N. crassa</i>]
AZM13183	MG06155	(MG06155.1) predicted protein
AZM13075	MG02288	(MG02288.1) predicted protein
AZM13031	MG02257	hypothetical protein [<i>N. crassa</i>]
AZM10813	MG10720	ribonuclease F1 [<i>Gibberella fujikuroi</i>]
AZM10789	MG10699	hypothetical protein [<i>N. crassa</i>]
AZM10800	MG10708	(MG10708.1) predicted protein
AZM10735	MG10679	(MG10679.1) predicted protein
AZM11007	MG10865	(MG10865.1) predicted protein
AZM11023	MG10876	hypothetical protein [<i>N. crassa</i>]
AZM10957	MG10824	(MG10824.1) predicted protein
AZM10372	MG10423	glycosyl hydrolase family 5/cellulase; At3g26140.1 [<i>A. thaliana</i>]
AZM10458	MG10471	(MG10471.1) predicted protein
AZM10421	MG10460	(MG10460.1) predicted protein
AZM10615	MG10586	predicted protein [<i>N. crassa</i>]
AZM10512	MG05954	(MG05954.1) hypothetical protein
AZM10474	MG10484	(MG10484.1) predicted protein
AZM11535	MG01275	(MG01275.1) predicted protein
AZM11874	MG01495	(MG01495.1) predicted protein
AZM11787	MG01444	(MG01444.1) predicted protein
AZM02333	MG04237	(MG04237.1) predicted protein
AZM02456	MG00374	predicted protein [<i>N. crassa</i>]
AZM01843	MG03891	(MG03891.1) predicted protein
AZM01918	MG06327	aspartic protease [<i>A. oryzae</i>]
AZM02009	MG04015	related to DFG5 protein [<i>N. crassa</i>]
AZM01992	MG04001	related to MID1 protein [<i>N. crassa</i>]
AZM03126	MG04737	(MG04737.1) predicted protein
AZM03012	MG04689	hypothetical protein [<i>N. crassa</i>]
AZM03344	MG04880	related to spore coat protein SP96 precursor [<i>N. crassa</i>]
AZM03293	MG04841	(MG04841.1) predicted protein
AZM02748	MG04541	(MG04541.1) predicted protein
AZM02823	MG04583	(MG04583.1) predicted protein
AZM02803	MG04572	(MG04572.1) predicted protein
AZM02843	MG04599	(MG04599.1) predicted protein
AZM00566	MG03043	(MG03043.1) hypothetical protein
AZM00552	MG03042	(MG03042.1) predicted protein
AZM00818	MG03208	related to beta (1-3) glucanoyltransferase [<i>N. crassa</i>]
AZM00127	MG02756	(MG02756.1) predicted protein
AZM00082	MG02720	(MG02720.1) predicted protein
AZM00263	MG02849	aspartic protease [<i>A. oryzae</i>]
AZM00313	MG02884	putative membrane-associated or secreted protein [<i>M. grisea</i>]
AZM01347	MG03575	(MG03575.1) predicted protein
AZM01627	MG00319	Endoglucanase EG-II (Endo-1,4-beta-glucanase) [<i>T. reesei</i>]
AZM01545	MG03691	(MG03691.1) predicted protein
AZM01047	MG03368	(MG03368.1) predicted protein

Table 3.5 (continued)

AZM01027	MG03354	(MG03354.1) predicted protein
AZM01115	MG03417	(MG03417.1) predicted protein
AZM01158	MG03442	predicted protein [<i>N. crassa</i>]
AZM05701	MG05638	(MG05638.1) predicted protein
AZM05668	MG05635	(MG05635.1) predicted protein
AZM05567	MG07331	probable GEL1 protein [<i>N. crassa</i>]
AZM05635	MG05632	(MG05632.1) predicted protein
AZM05619	MG07362	(MG07362.1) predicted protein
AZM05890	MG00047	(MG00047.1) predicted protein
AZM05980	MG07598	(MG07598.1) predicted protein
AZM05292	MG07172	(MG07172.1) predicted protein
AZM05306	MG07179	Chain A, Pepsin Hydrolase (Acid Proteinase) (E.C.3.4.23.1)
AZM05189	MG07100	(MG07100.1) predicted protein
AZM05470	MG07264	glutaminase A [<i>A. oryzae</i>]
AZM05531	MG07306	related to mixed-linked glucanase precursor MLG1 [<i>N. crassa</i>]
AZM05507	MG07287	lysophospholipase [<i>N. crassa</i>]
AZM05379	MG00742	(MG00742.1) predicted protein
AZM05385	MG00737	(MG00737.1) predicted protein
AZM06557	MG07975	(MG07975.1) predicted protein
AZM06616	MG05705	(MG05705.1) predicted protein
AZM06609	MG08014	hypothetical protein [<i>N. crassa</i>]
AZM06793	MG05456	(MG05456.1) predicted protein
AZM06401	MG00815	(MG00815.1) predicted protein
AZM06301	MG07807	(MG07807.1) predicted protein
AZM06261	MG07775	(MG07775.1) predicted protein
AZM03988	MG05270	(MG05270.1) predicted protein
AZM03856	MG00539	(MG00539.1) predicted protein
AZM04203	MG05393	(MG05393.1) predicted protein
AZM04222	MG05406	(MG05406.1) predicted protein
AZM04142	MG00592	Crh-like protein [<i>A. fumigatus</i>]
AZM03543	MG04991	hypothetical protein [<i>N. crassa</i>]
AZM03755	MG05120	(MG05120.1) predicted protein
AZM03843	MG00552	acid phosphatase PHOa [<i>A. fumigatus</i>]
AZM03814	MG05165	(MG05165.1) predicted protein
AZM03660	MG05069	(MG05069.1) predicted protein
AZM03681	MG05083	(MG05083.1) predicted protein
AZM03732	MG00527	(MG00527.1) predicted protein
AZM05053	MG07005	(MG07005.1) predicted protein
AZM05014	MG00721	(MG00721.1) predicted protein
AZM04609	MG06722	probable beta (1-3) glucanosyltransferase [<i>N. crassa</i>]
AZM04519	MG06663	endo-1,4-beta-glucanase [<i>A. oryzae</i>]
AZM04533	MG06653	hypothetical protein [<i>N. crassa</i>]

^a Agilent probe ID

^b *Magnapothe grisea* gene ID

Figure legends

Figure 3.1 Sample preparation for microarray analysis.

Microsome-bound RNA (MBP) and free cytosolic RNA (FP) were prepared using both differential centrifugation and sucrose gradient centrifugation. Microarray analysis was performed using differential centrifugation preparation as described in the materials and methods. The microarray data was confirmed using samples prepared by both methods for MM48. RT-PCR analysis was performed using differential centrifugation preparation for APP.

Figure 3.2 Partitioning of the genes in ergosterol biosynthesis pathway.

Enzymes in ergosterol biosynthetic pathways are listed in the figure. The data was taken from appressorial microarray experiment. The locations of the enzymes are from GO cellular component. Compared to the early steps of ergosterol biosynthetic steps, late steps are mainly suggested to be associated with the membrane. Enrichment level is shown in red-blue scale. Red colors indicate transcripts predominantly in MBP, and blue colors represent transcripts that are predominantly found in FP.

Figure 3.3 The distribution of MBP:FP signal ratio for transcripts encoding proteins with signal peptide.

(A) All genes containing a predicted signal peptide. The graph indicates that the majority of genes with a predicted signal peptide exhibited MBP:FP ratio >1 . (B) Genes predicted by Dean et al. (2005) to be secreted. Similar results to (A) are observed.

Figure 3.4 RT-PCR analysis of microarray result.

Semi-quantitative RT-PCR was done to confirm microarray data. The presence and absence of predicted signal peptide and enrichment of the transcripts in MBP or FP are indicated. (A) RT-PCR analysis of MM48 samples. RNA samples were prepared using differential centrifugation method (left panel) and sucrose gradient method (right panel). The results show a similar trend. (B) RT-PCR analysis of APP sample. RNA sample was prepared using differential centrifugation method. The MG identifications, their corresponding probe IDs, and description of the genes are as follows: MG06432 (AZM03364) predicted protein, MG06594 (AZM04426) extracellular chitinase, MG02503 (AZM13403) heat shock 70kD protein 5 (BiP), MG09736 (AZM09326) predicted protein, MG07975(AZM06557) predicted protein, MG01931 (AZM12495) MFS efflux transporter of unknown specificity in *S. pombe*, MG3416 (AZM01114) predicted protein, MG10315 (AZM10225) MPG1 hydrophobin, MG5110 (AZM03716), MG08264 (AZM06992) predicted protein, MG09082 (AZM08292) predicted protein with homology in *N. crassa*, MG10660 (AZM10704) predicted transcription factor, MG08164 (AZM06847) probable protein disulfide-isomerase precursor in *N. crassa*, MG01472 (AZM11841) ArfGAP(GTPase-activating protein) in *S. cerevisiae*. (C) MG04582 β -glucosidase (AZM02821) is differentially enriched in MBP in APP and FP in MM48.

Figure 3.5 Comparison of hydropathy plot of signal peptides.

Hydropathy plots were generated using the signal peptides from 3 transcripts that are enriched in each fraction were taken for signalP analysis. To generate the hydropathy plot, EMBOSS package was used using pepwindow program using the default setting as

described in the methods. The hydropathy plots suggest that the enrichment may not be due to the strength of signal peptide (hydrophobicity) as reported in *S. cerevisiae*.

Figure 3.6 Translational activities of select *M. grisea* transcripts.

Total RNA from MM48 was subject to polysomal separation using a sucrose gradient. Panel (A) shows the MM48 polysomal profile. Panel (B) is Northern analysis of select transcripts. The RNA samples are loaded in the order of free to polysome-bound RNAs (lane 1~6) and separated on denaturing gel, and Northern analysis was performed using the probes indicated. MPG1 transcripts are enriched in the FP fraction for both MM48 and APP conditions. In contrast, BiP is enriched in the MBP fraction in both condition. MG10660, a putative transcription factor, lacking a signal peptide, was used as a control for FP enriched transcripts. The result indicates that all three transcripts are associated with ribosomes to a varying degree, and translationally active.

Figure 3.7 Differential staining of chitin and chitosan during infection.

Appressorium formation was induced on cellophane film (A~D) and onion skin (E~H) for staining of chitin and chitosan. Incipient penetration peg was stained with WGA-TRICT for chitin (B and F) and chitosan-specific antiserum was used to stain chitosan (A and E) followed by incubation with anti-rabbit FITC-conjugated antibody. The composite image and bright field image are shown in C and D for cellophane film and G and H on onion skin, respectively. Compared to the staining of germ tubes and conidium, infection hyphae showed more intense staining with chitosan-specific antibody while less intense staining was observed for chitin by WGA. (I) RT-PCR analysis of chitin deacetylase suggested that the expression of chitin deacetylase was sensitive to nutrient, and slight up

regulation was also noticed during the late stage of infection 24 hours and 48 hours after incubation (MM: minimal medium, -N: MM-N, -C, MM-C, A7: appressorium after 7 hours of induction, A24: appressorium 24 hours of induction, A48: appressorium 48 hours of induction).

Figure 3.8 The distribution of MBP:FP signal ratio for MAGGY components (gag and pol) grown for 48 hrs in minimal medium.

The ratio of MBP:FP is shown in a log scale. The distribution of MBP:FP ratio indicates that the retrotransposon, MAGGY, is associated with (endo)membrane fraction in *M. grisea*.

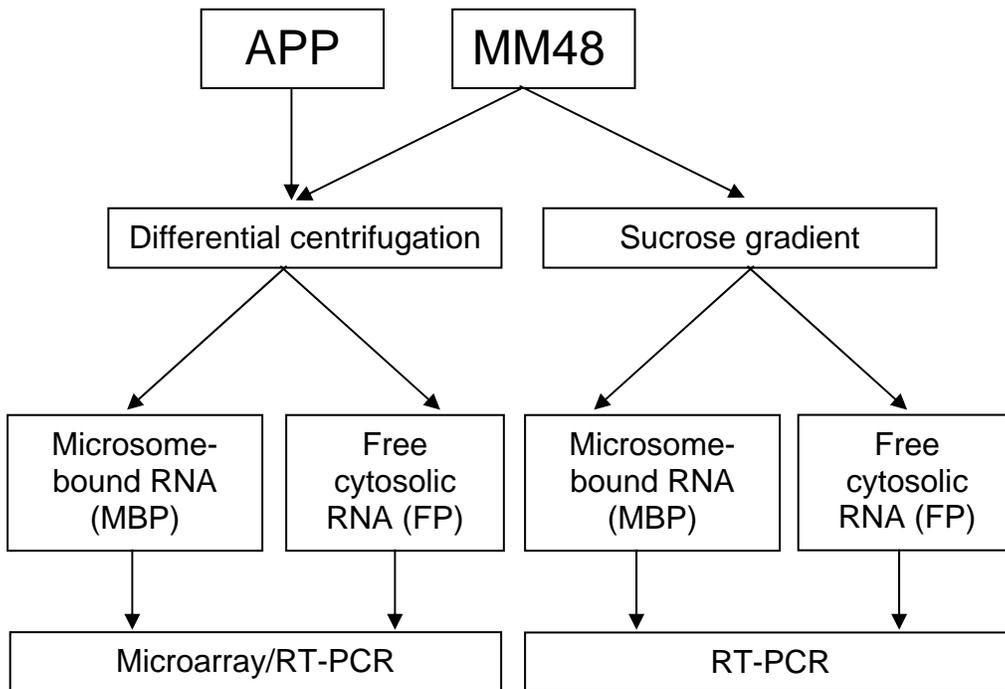


Figure 3.1 Sample preparation for microarray analysis

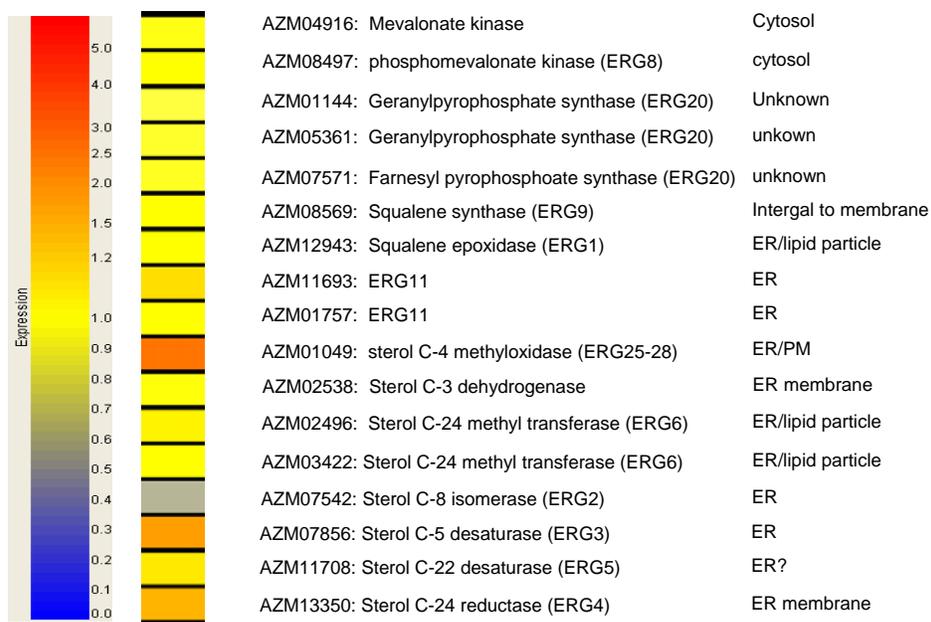
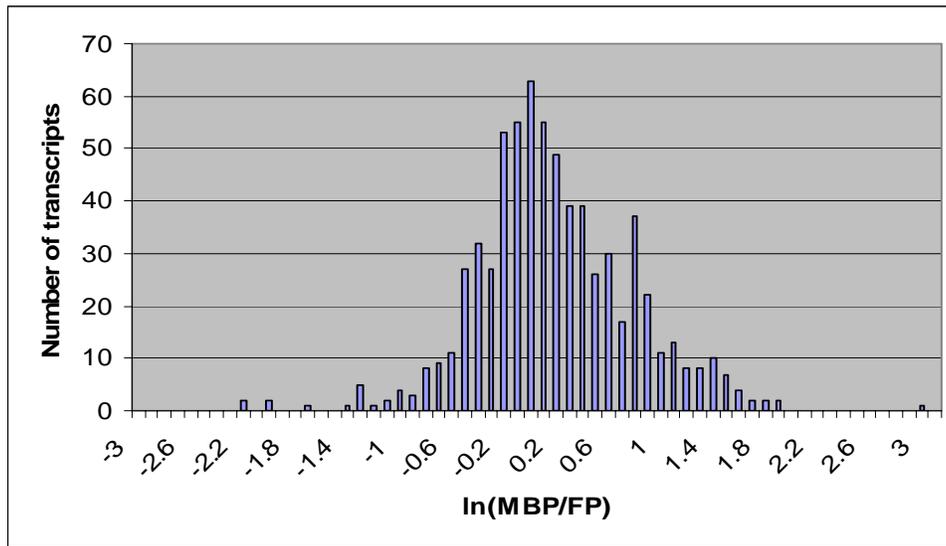


Figure 3.2 Partitioning of the genes in ergosterol biosynthesis pathway

(A) Genes with predicted signal peptides



(B) Published secretome

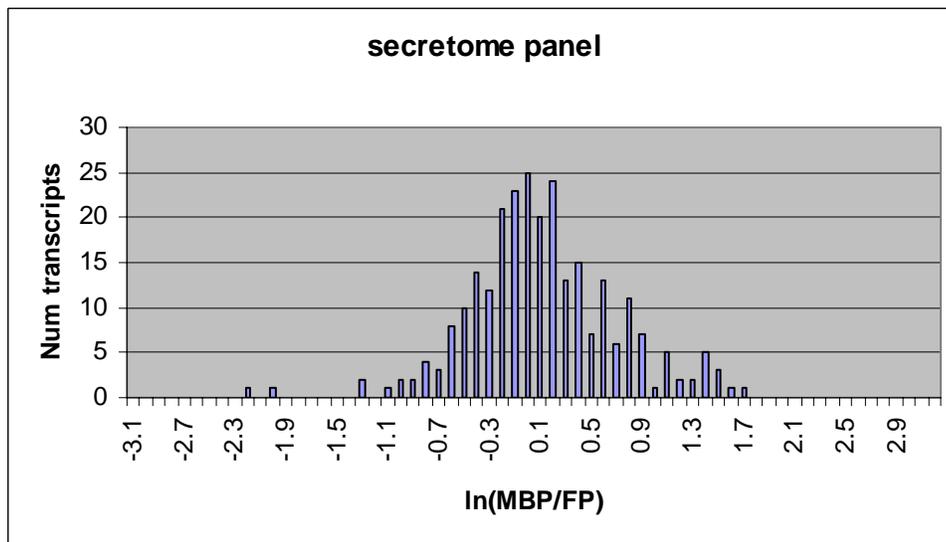


Figure 3.3 The distribution of MBP:FP signal ratio for transcripts encoding proteins with signal peptide

(A)

Gene	Differential		Density		Signal pep	Enrichment
	FP	MBP	FP	MBP		
MG06432					-	MBP
MG06594					+	FP
MG02503					+	MBP
MG09736					+	FP
MG07975					-	FP
MG01931					-	MBP
MG03416					-	FP
MG10315					+	FP

(B)

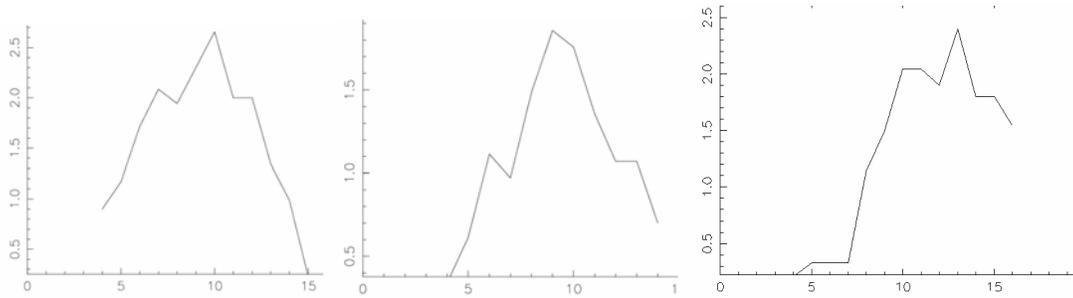
MG05110			-	MBP
MG08264			+	FP
MG09082			-	MBP
MG10315			+	FP
MG10660			-	FP
MG08164			+	MBP
MG01472			-	MBP
MG02647			+	MBP

Figure 3.4 (continued)

(C)



Figure 3.4 RT-PCR analysis of microarray result

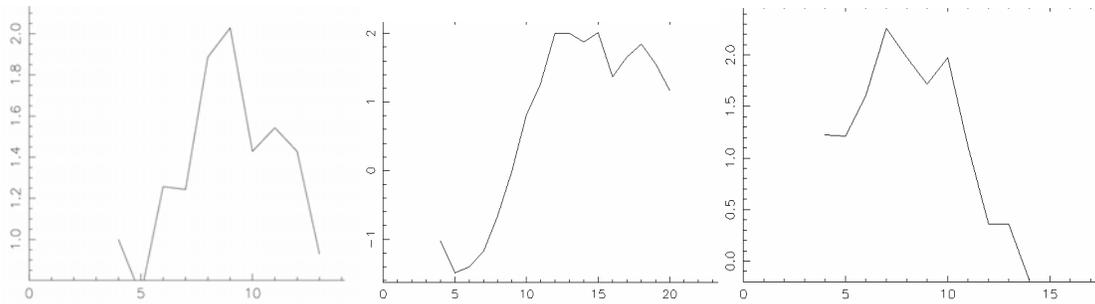


<MG08480>

<MG08817>

<MG09385>

FP-enriched



<MG05689>

<MG05515>

<MG07153>

MBP-enriched

Figure 3.5 Comparison of hydropathy plot of signal peptides

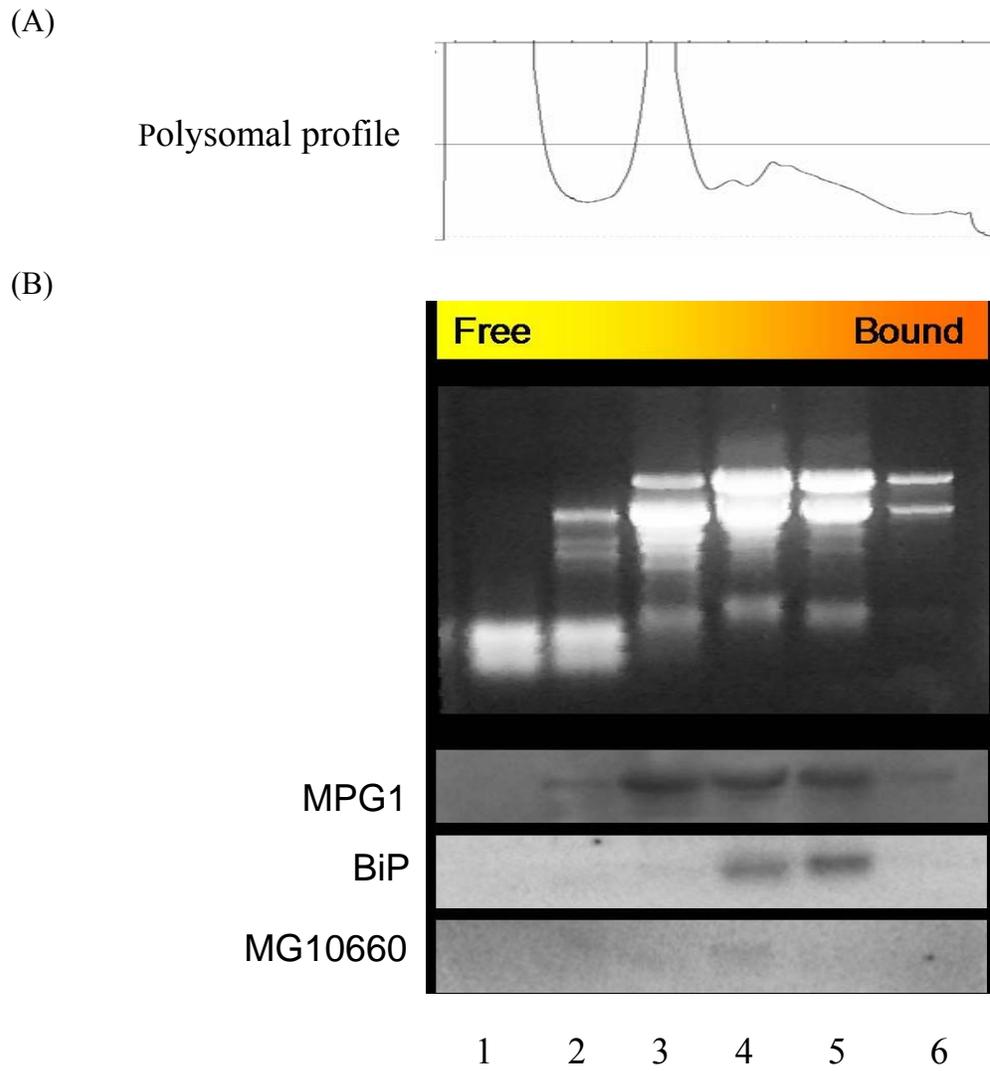
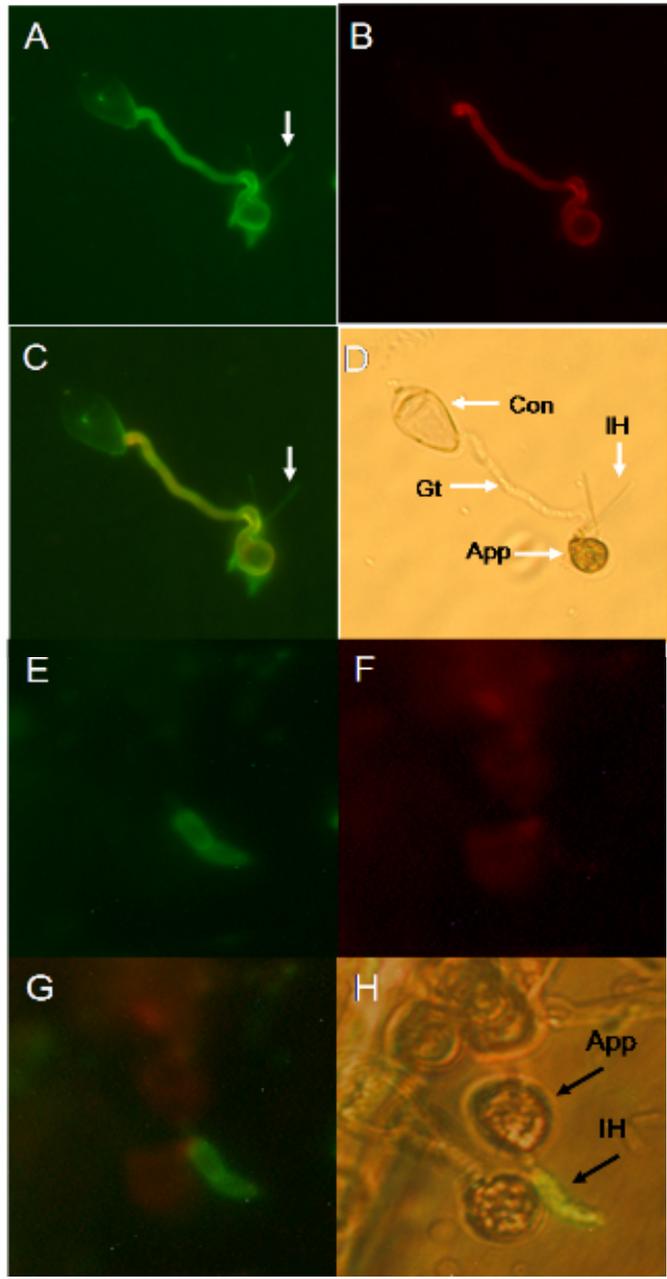


Figure 3.6 Translational activities of select *M. grisea* transcripts



I

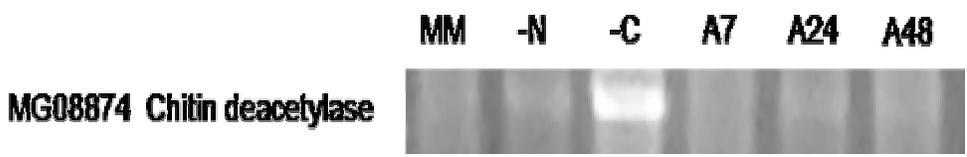


Figure 3.7 Differential staining of chitin and chitosan during infection

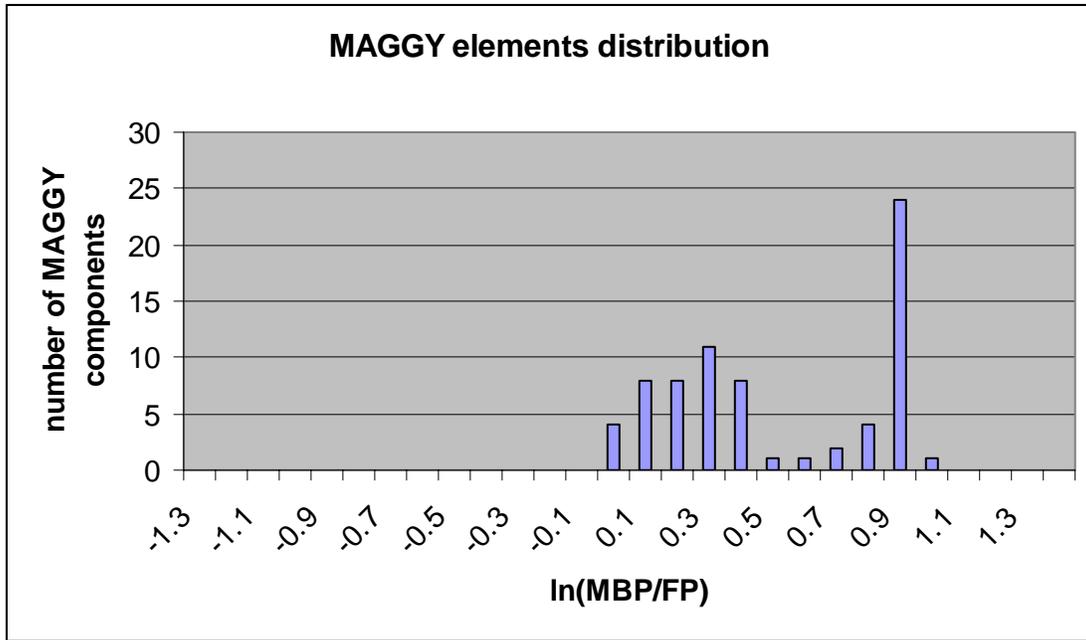


Figure 3.8 The distribution of MBP:FP signal ratio for MAGGY components (gag and pol) grown for 48 hrs in minimal medium

Chapter Four

14-3-3 homolog, MgFTT1, in *Magnaporthe grisea* affects carbon metabolism and is required for growth *in planta*

Jun Seop Jeong and Ralph A. Dean

Center for Integrated Fungal Research (CIFR), North Carolina State University,
Raleigh, NC 27695

Abstract

14-3-3 is a ubiquitously found protein family in eukaryotes. Despite extensive research in a variety of organisms, definitive roles for 14-3-3 remain elusive. In yeast-like fungi, 14-3-3 proteins are known to regulate several aspects of growth and metabolism including polarized growth and nutrient sensing. Studies of 14-3-3 in many fungi have been hampered by the inability to recover gene knockout mutants and/or apparent functional redundancy. In this work, we characterized a 14-3-3 homolog in the rice blast fungus, *Magnaporthe grisea*. Two copies of 14-3-3 were identified in the genome, and one copy, MgFTT1, was functionally evaluated. Deletion strains of MgFTT1 created through double homologous recombination exhibited alterations in extracellular carbon-sensitive enzymatic activities. Mutants also exhibited reduced pathogenicity to host plants which in part were due to an impaired ability to colonize host tissue. Localization studies using a translational fusion of MgFTT1 with GFP revealed nuclear cytoplasmic shuttling under different conditions, indicative of dynamic subcellular localization.

Introduction

Rice blast is the most important disease in rice production world-wide. The disease is caused by the ascomycetous filamentous fungus *Magnaporthe grisea*. The genome of the rice blast fungus was recently sequenced, revealing a wealth of information regarding potential pathogenicity genes (Dean et al., 2005). For example, the presence of batteries of hydrolytic enzymes such as cutinases suggests a genomic adaptation to its ecological niche as a pathogen compared to the saprophytic fungus *Neurospora crassa* (Galagan et al., 2003). The infection process of the rice blast fungus is initiated by the formation of anisotropic structures, called appressoria. Appressoria develop from the narrow germ tube following conidial germination on the leaf surface. Appressorium development is accompanied by the generation of hydrostatic turgor pressure resulting from glycerol accumulation (de Jong et al., 1997). The turgor pressure enables a penetration peg to breach the host surface mechanically.

Mobilization of internal storage molecules for the generation of glycerol appears to be the result of several signal transduction pathways (Dixon et al., 1999). cAMP signaling and MAP kinase pathways have been extensively studied in *M. grisea*. Application of exogenous cAMP and its analogs were shown to induce the appressorium development (Lee and Dean, 1993). The main target of cAMP is the protein kinase A (PKA). Deletion of PKA catalytic subunit results in abnormal appressorium formation that fail to penetrate the host surface, resulting in a complete loss of pathogenicity (Mitchell and Dean, 1995; Xu et al., 1997). A MAC1 (adenylate cyclase) mutant showed pleiotropic effects in development and loss of pathogenicity, corroborating the role of cAMP in pathogenicity (Choi and Dean, 1997). The MAP kinase cascade is a canonical

phosphorylation cascade with three kinases in distinct modules responding to different environmental signals in eukaryotes (Garrington and Johnson, 1999). For example, loss of function of FUS/KSS1-related MAP kinase results in defects in appressorium formation and infectious growth in *M. grisea* (Xu and Hamer, 1996). The MAP kinase pathway has been shown to have a fundamental importance in pathogenesis in *M. grisea* and other fungal pathogens (Xu, 2000).

14-3-3 is a conserved protein family found in all eukaryotes examined, and is a potential downstream target of signal pathways (Fu et al., 2000). However, despite intensive research on the role of 14-3-3 proteins in higher eukaryotes, their definitive function remains elusive. One of the well-established roles of 14-3-3 proteins is the binding of interaction partners and modulation of protein activity (Yaffe and Elia, 2001). The interaction is mainly mediated by binding to phosphorylated targets which participate in a variety of signal transduction pathways (Yaffe, 2002). 14-3-3 proteins bind a diverse array of targets including protein kinases, transcription factors, structural proteins, and enzymes (Yaffe, 2002). In *Saccharomyces cerevisiae*, the 14-3-3 protein homologs, BMH1 and BMH2, mediate signal transduction in the FUS3/KSS1 MAPK pathway, which participates in invasive growth and pseudohyphal growth (Roberts and Mosch, 1997). The binding of 14-3-3 proteins to interacting partners is known to modulate enzymatic activity or subcellular localization. 14-3-3 also plays an important role in transcriptional regulation (Beck and Hall, 1999; Igarashi et al., 2001). BMH1 and BMH2 bind to stress related transcription factors, and sequester them in the cytoplasm. In addition, 14-3-3 proteins function through the masking of important motifs in interacting partners such as phosphorylation sites, nuclear export signals (NES), and other regulatory

sequences (Lopez-Girona et al., 1999). Furthermore, 14-3-3 protein mediates intracellular trafficking (Gelperin et al., 1995; Roth et al., 1999). Despite the universal presence of 14-3-3 proteins in eukaryotes and its diverse role in cell biology, in most studies, the phenotypes of deletion of either BMH1 and BMH2 are subtle considering the fundamental and critical cellular processes the proteins are involved in (van Heusden et al., 1992; Roberts and Mosch, 1997). Functional redundancy of 14-3-3 proteins in a genome may account for the observations. However, characterization of 14-3-3 in yeast provided evidence that 14-3-3s are linked to nutrient utilization and signaling pathways (van Heusden et al., 1992; Bertram et al., 1998; Roberts and Mosch, 1997).

In addition to extensive characterization in *S. cerevisiae*, the function of 14-3-3 has been studied in some detail in other yeast-like fungi. 14-3-3 protein homologs in fission yeast were found to bind Byr2 (MAPKKK) (Ozoe et al., 2002), supporting its regulatory role in the MAP kinase pathway. A similar role was found in *Yarrowia lipolytica* (Hurtado and Rachubinski, 2002) and *Candida albicans* (Cognetti et al., 2002). Interestingly, 14-3-3 protein was identified in *Y. lipolytica* as a suppressor of a small GTPase Rac-mediated defect in filamentous growth (Hurtado and Rachubinski, 2002). GTPase modules such as heterotrimeric G-proteins, ras, Rho, Rac and cdc42 are located upstream of the protein kinase pathways including PKA and PKC, and MAPK pathways (Banuett, 1998).

Compared to work in yeast species, relatively little is known in filamentous fungi. In *Aspergillus nidulans*, 14-3-3 functions in polarized hyphal growth (Kraus et al., 2002). Other studies using *Trichoderma reesei* (Vasara et al., 2002) have suggested that the protein is involved in protein secretion, suggesting a conserved function in cytoskeleton

reorganization for protein secretion in as observed in yeast (Roth et al., 1999). Studies in filamentous fungi have been primarily limited to the controlled over-expression strategy due to difficulties with knockout of 14-3-3 gene, presumably resulting from the apparent essential nature of 14-3-3. In this study, we identified two 14-3-3 homologs in the *M. grisea* genome, and successfully deleted one, MgFTT1. Phenotypic analyses suggested that the protein plays a regulatory role in response to stress. The capacity of deletion mutants to form appressorium was not altered, and the germinating conidia were responsive to inducers, including hydrophobicity, cutin monomer, and cAMP regulators. However, the deletion of the gene did affect the ability of the fungus to colonize susceptible host tissue, suggesting its importance as a general regulatory factor. Investigation of intracellular localization using a GFP fusion suggested a dynamic subcellular localization.

Materials and methods

Fungal strains and culture condition

Magnaporthe grisea 70-15 was recently sequenced (Dean et al., 2005), and used for this study. 70-15 was grown on oatmeal agar (OMA) medium for routine propagation under constant light. For nucleic acids preparation, the culture was grown in liquid complete media (CM; 6g yeast extracts, 6 g casein, 10 g sucrose per liter). The composition of minimal medium (MM) was as follows: sucrose 10 g, *A. nidulans* trace element solution 1 ml, 20x nitrate salts 50 mL, thiamine 1 mg, and biotin 5 ug per liter. The trace element solution was NaNO₃ 60 g, KCl 5.2 g, MgSO₄7H₂O 5.2 g, KH₂PO₄ 15.2

g in 500 ml. For the preparation of MM-C and MM-N, sucrose and NaNO₃ were omitted, respectively.

Identification of MgFTT in the genome

One 14-3-3 protein domain containing open reading frame was annotated in Broad Institute database (www.broad.mit.edu/annotation/fungi/Magnaporthe/) which we named MgFTT1. 14-3-3 homologs from fungal species were retrieved from RefSeq GenBank database. BLAST analysis (Altschul et al., 1997) was done to identify an additional 14-3-3 paralog in the rice blast fungus genome. Sequence alignment was done using ClustalW (Chenna et al., 2003) software and phylogenetic analysis was performed using MEGA software (version 3) (Kumar et al., 2004). The alignment was carefully inspected, and the phylogenetic tree was generated using the maximum parsimony method. Bootstrap was performed for 1000 iterations.

Generation of Knockout mutants

For generation of deletion mutants, sequence flanking MgFTT1 was amplified using 1433speI primer (GAAGAACTAGTCCTGGGAGCTACTAATTTCG) and 1433-5ORFout primer (GGCCGGCCGAATTCGTAAACGGCATCTTCGTGACCCAT) for the flanking sequence on the left side and 1433sal (GAAGAAGTCGACGGCACCTAACTTATCTTGGC) and 3'ORFout (CACGGCGCGCCTAGCAGCGGAATTCCTGGTATCGATGGTGGATTG) primers for the cloning of the right border. The left flanking sequence was ligated to pCB1179 which contains a copy of the hygromycin gene (Carroll

et al., 1994) using SpeI and EcoRI sites, generating pFTT1-1. The right flanking sequence was ligated into pFTT1-1 using end-polished SacII site after blunt-ending of the PCR product, generating pFTT1-2. The vector was linearized using SpeI restriction enzyme before transformation. Approximately, 5 µg of DNA was used for each transformation using the PEG-mediated transformation method (Leung et al., 1990). Hygromycin resistant colonies were selected on TB3 agar amended with 400 µg/ml hygromycin. The outgrowing colonies were transferred to complete media amended with 400 µg/ml of hygromycin for secondary selection. Colonies with confirmed growth on antibiotic media were screened for deletion events by PCR. For detection of the deletion event, primers 14-3-3NcoF (GAAGGAGATAGAACCATGGGTCACGAAGATGCCG) and 14-3-3EcoR (CAAGAAAGCTGGGTCGAATTCCTCAGCAGGCTTCGG) primers were used. The confirmation of the deletion event was done using genomic DNA Southern hybridization using PCR product from 14-3-3NcoF and 14-3-3EcoR primers as a probe. For the cloning of the MgFTT1 ORF into the GATEWAY cloning system, the ORF PCR product was amplified with attB1univ (GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAGGAGATAGAAACCATG) and attB2univ (GGGGACCACTTTGTACAAGAAAGCTGGGT) primers. The expression vector pLD2 was a kind gift from Dan Ebbole at Texas A&M University. Expression is directed under the PR27 promoter, and the construct harbors the RGSHHHHHHH epitope and the GFP coding region. The vector was modified to be compatible with the GATEWAY cloning system by inserting the GATEWAY vector conversion cassette into the SmaI site in frame according to the manufacturer's instruction. The vector was designated pLD2-GW. MgFTT1 ORF was cloned in a donor vector pDONR221 using the BP reaction and the insert was

subsequently transferred to pLD2-GW using the LR reaction. The resulting construct, pLD2-FTT1 was used for MgFTT1::GFP fusion overexpression and localization studies.

Microscopic studies

To determine the subcellular localization of the protein *in vivo*, the GFP fusion construct was transformed into the wild type strain. GFP expression was observed using Zeiss fluorescence microscope using the FITC filter. Microscopic observation of glycogen body was made by staining cells with Lugol solution, and Nile red was used to stain the lipid body, and observation was made using the FITC filter set (Both et al., 2005). For fluorescence microscopy, wheat germ agglutinin (WGA)-TRITC or Calcofluor were used to stain chitin and nuclei respectively, and observation was made using texas-red and DAPI filters.

Appressorium assays

Appressorium assays using GelBond and inducers were performed as described previously (Lee and Dean, 1993). Appressorium induction using cutin monomers followed published methods (Gilbert et al., 1996). The cellophane penetration assay was done as follows: Fifty microliters of spore suspension (10^5 conidia/ml) was spotted on sterilized cellophane film placed on OMA. After incubation for 48 hrs, the cellophane film was removed, and the OMA plate was further incubated until visible fungal growth was seen. A *pth11* deletion strain was generated by J. Marui, FGL, and was a kind gift

for this research. The phenotype of pth11 is described elsewhere (DeZwaan et al., 1999). Each experiment was repeated at least three times with three replicates.

Enzymatic assay

The invertase assay was done following an established method using chromogenic *o*-dianisine (Darsow, Odorizzi, et al. 2000 113 /id). The activity of invertase was measured spectrophotometrically by reading O.D. at 540 nm. β -Galactosidase and β -glucosidase activity were measured using *o*-nitrophenyl- β -D-galactopyranoside (ONPG) (Ausubel et al., 1987) and *p*-nitrophenyl β -D-glucopyranoside (Lee et al., 1996), respectively, at a wavelength of 410 nm. For enzyme assays, conidial suspension was inoculated in MM liquid medium. The culture was grown with with gentle shaking for 24 hrs at room temperature. The culture filtrate was collected and assayed for invertase, β -galactosidase, and β -glucosidase. The experiment was repeated three times with five replicates, similar pattern of color development was observed in all experiments.

Pathogenicity assay

Conidial suspension was collected from cultures grown on OMA plates, adjusting to 10^5 conidia/ml, and was spray inoculated onto 2 weeks old rice seedling of the susceptible cultivar Maratelli with five replicates. The progress of the infection was monitored through 2 weeks. Pathogenicity was also tested on barley in a similar manner. For wound assays, two weeks old detached barley seedlings were pricked with sterile pipette tips, and 50 μ l of 10^4 conidia/ml suspensions were inoculated on the wounded

sites. The progression of the symptoms was photographed 1 week after inoculation. Tests were repeated at least three times.

Results

Identification of MgFTT1

A 14-3-3 homolog (MG01588) was identified in *M. grisea* genome and named MgFTT1. BLASTP search using the MgFTT1 protein sequence retrieved homologs of 14-3-3 proteins in other fungi with the closest homologs found in *Hypocrea jecorina* (CAC20377.1: 96% identity), 14-3-3TRIHA 14-3-3 protein homolog in *Gibberella zeae* (EAA76369.1: 96% identity), and 14-3-3 protein homolog in *Neurospora crassa* (XP_329994.1: 93% identity). BLAST analysis of the rice blast genome and TIGR Index EST database revealed the presence of another copy of 14-3-3 gene, named MgFTT2. The presence of this gene was confirmed by RT-PCR. Pair-wise sequence alignment of MgFTT1 and MgFTT2 revealed in 69% identity and 83% similarity by PAM250. A phylogenetic tree generated using maximum parsimony method using available 14-3-3 fungal sequences indicated a clear separation of 14-3-3 sequences between Ascomycetes and Basidiomycetes (Figure 4.1). The 14-3-3 homologs from *S. cerevisiae* BMH1 and BMH2 formed a distinct branch from other Ascomycetes. The results also clearly revealed two distinct isoforms of 14-3-3 proteins in filamentous fungi. For example, MgFTT1 and MgFTT2 were located in separate branches, suggesting isoform-specific function.

Expression analysis of MgFTT1

The expression pattern of MgFTT1 was analyzed by Northern analysis (Figure 4.2). RNA samples were prepared from conidia, germinating conidia, and mycelia grown in CM, MM, and MM without either carbon (MM-C) or nitrogen sources (MM-N). Transcripts analysis showed that expression of MgFTT1 was detectable under all conditions tested. However, transcript levels were higher in germlings compared to conidia and the appressorial stage. We also observed fluctuations in transcripts level in mycelia grown in rich versus minimal media and in response to nutritional stress, suggesting MgFTT1 might be linked to polar growth and nutrient control. This is consistent with proposed roles of the protein in other fungal species.

Generation of knockout mutants

Genomes of higher eukaryotes typically contain many copies of 14-3-3 homologs. Fungi, on the other hand, typically contain 2 copies. Deletion of either BMH1 or BMH2, 14-3-3 homologs in *S. cerevisiae*, is not lethal. However, in most genetic backgrounds, the double mutation is lethal. In other yeast species such as *C. albicans* and *Y. lipolytica*, 14-3-3 appears to be essential for viability (Cognetti et al., 2002; Hurtado and Rachubinski, 2002). *C. albicans* only appears to contain a single copy of 14-3-3 whereas *Y. lipolytica* contains two copies of 14-3-3 homologs. However, attempted deletion of a 14-3-3 homolog, YIBMH1 was not successful in *Y. lipolytica*, suggesting that this gene is essential for viability. In *T. reesei*, the function of the gene was characterized by complementation of a truncated cDNA fragment in budding yeast for a secretion defect (Vasara et al., 2002). Work in *A. nidulans* also suggested that the 14-3-3 homolog, ArtA,

may be essential for viability (Kraus et al., 2002). To further the knowledge of 14-3-3 in filamentous fungi, we attempted to delete MgFTT1 from the *M. grisea* genome. Three independent deletion mutants were obtained by the homologous recombination strategy depicted in Figure 4.3A. To confirm deletion of the gene, genomic Southern hybridization analysis was performed using the ORF as a probe (Figure 4.3B upper panel). The hybridization data showed the absence of the signal in Δ MgFTT1-1, 2 and 3, and hybridization with hph probe showed integration of hph in all strains except for wild type (Figure 4.3B lower panel). PCR analysis confirmed the deletion of MgFTT1 from strains Δ MgFTT1-1, 2, and 3. The growth of deletion strains on OMA was comparable to wild type in terms of growth rate and pigmentation. Conidiation level appeared to be indistinguishable from the wild type with the exception of Δ MgFTT1-3 which appeared to have a slightly reduced ability to conidiate.

Deletion of 14-3-3 affects the expression of hydrolytic enzymes

14-3-3 is known to regulate starvation responses in yeast (Bertram et al., 1998; Beck and Hall, 1999) . In yeast, BMH1 and BMH2 binds to the stress responsive transcription factor MSN2 and MSN4, sequesters them in the cytoplasm, and only releases the transcription factors upon a nutrient deprivation signal (Beck and Hall, 1999). The two transcription factors are known to be activated by stress such as nutrient starvation. Invertase activity is affected by the expression of *T. reesei* 14-3-3 homolog in yeast (Vasara et al., 2002). To see if the deletion of MgFTT1 affects the expression of the enzymatic activities that may be under catabolite repression, enzymatic activity was measured in the culture filtrates. We choose invertase, β -galactosidase, and β -glucosidase

activities. In yeast, invertase and β -galactosidase are repressed in the presence of good carbon sources (Gancedo, 1998). Also, β -glucosidase activity has been shown to be regulated by carbon catabolite repression in filamentous fungi (Lee et al., 1996). The results are summarized in Figure 4.4. Knockout of MgFTT1 resulted in a 5-10 fold increase in invertase activity in the culture filtrate compared to the wild type and ectopic mutant (Figure 4.4A). Similar effects were seen with the β -galactosidase (Figure 4.4B) and β -glucosidase (Figure 4.4C) activities. Thus, the data indicate that MgFTT1 regulate a number of extracellular hydrolases, potentially through a nutrient sensitive mechanism.

Deletion of MgFTT1 does not affect the ability to form appressorium or to penetrate

One of the proposed roles of 14-3-3 in a variety of fungi is the regulation of polarized growth. Also, BMH1 and BMH2 in yeast play a central role in the RAS/cAMP pathway. In the rice blast fungus, appressorium formation is essential for infection, and is known to be mediated by cAMP. Thus, we investigated the effect of MgFTT1 deletion on appressorium formation in *M. grisea*. Deletion mutants retained the capacity to form appressoria on a hydrophobic surface. Induction of appressoria with cAMP, a cyclic nucleotide phosphodiesterase inhibitor, IBMX, and a cutin monomer 1,16-hexadecanediol induced appressorium on non-inductive surface indistinguishable from wild type. Staining of lipid body and glycogen by Nile red and Lugol solution did not show difference between wild type and the deletion strains. Thus, the ability to form appressorium does not appear to be affected by deletion of MgFTT1.

Genetic and physical interaction of 14-3-3 with Rac1 and PAKs has been reported in other fungi (Hurtado and Rachubinski, 2002; Szabo, 2001; Leveleki et al., 2004). In

Ustilago maydis, deletion of a PAK kinase resulted in alteration of chitin metabolism. Thus, we were interested to determine whether deletion of MgFTT1 altered the cell wall of *M. grisea*. Staining of germinating *M. grisea* conidia with a lectin, wheat germ agglutinin (WGA) TRICT, conjugate showed differences (Figure 4.5A). In the wild type strain, the cell walls of germ tubes and incipient appressorium were intensely stained. However, mutant strains showed a reduced level of staining. To see if the appressorium formed by the deletion mutants was defective in some way, the ability to penetrate an artificial substratum was tested (Figure 4.5B). An agar block containing fungal material was placed face-down on sterile cellophane film overlaid on OMA, and incubated for 48 hrs. After incubation, the cellophane was removed and the plates were further incubated until clear colony growth was visible. Clear growth was seen for all of the deletion strains, indicating penetration of the cellophane film. However, no growth for *pmk1*, *cpka*, *magb*, and *mac1* mutants was observed. However, growth of *pth11* deletion mutant was seen, suggesting that the penetration of cellophane film is strongly correlated to the formation of a functional appressorium and invasive growth. Thus, the data suggest that MgFTT1 deletion mutants form functional appressorium. We further investigated the ability of MgFTT1 deletion mutants to penetrate onion peel epidermal cells. We were able to observe infectious hyphae when inoculated with the deletion mutants (data not shown).

MgFTT1 is required for growth in planta

To investigate whether the deletion of the gene may affect the ability to cause disease, pathogenicity assays were done on rice and barley. Pathogenicity was assessed using two week old rice seedlings. Symptoms were evaluated two weeks after inoculation.

Even with an apparent absence of defects in appressorium formation and penetration, differences in pathogenicity was noticed in the deletion strains compared to the wild type. Wild type and the ectopic strain produced disease lesions that spread rapidly. However, the deletion mutants exhibited reduced symptom development as evidenced by the formation of fewer numbers and smaller lesions compared to the wild type (Figure 4.6A for rice and 4.6B for barley). To investigate whether the deletion strains were defective in invasion of host tissues, detached barley leaves were inoculated after wounding. In contrast to the lesions produced by the wild type and the ectopic mutant, which showed spreading lesions, the deletion strains exhibited limited symptom development on wounded barley leaves (Figure 4.6C). Thus, the data suggests that MgFTT1 is required for normal growth *in planta*.

14-3-3 GFP fusion protein localizes to the nucleus

PSORTII predicted a subcellular localization for MgFTT1, and suggested nuclear localization (Nakai and Horton, 1999). To investigate this, GFP was fused to the C-terminal of full length MgFTT1 under the constitutive RP27 promoter (Bourett et al., 2002). In *Arabidopsis*, a similar translational fusion of 14-3-3s exhibited a distinct subcellular localization, suggesting that translational fusion of GFP was functional (Paul et al., 2005). In *M. grisea*, the GFP signal was observed throughout the cytoplasm when the culture was grown in nutrient rich liquid medium. However, when transferred to CM amended with hygromycin, the GFP signal localized to the nucleus (Figure 4.7A and Figure 4.7B). The strain that contain multi-copy of the 14-3-3::GFP fusion showed some interesting phenotypes. Colony morphology and the growth characteristics were not

distinguishable in terms of growth rate and conidiation compared to the wild type on solid media. However, no development of pigmentation was seen in liquid culture of the overexpression strain whereas wild type cultures became pigmented after a prolonged growth. Strong subcellular localization was also seen in conidia. In fresh conidia, the GFP signal was absent from the tapered end of a conidium (Figure 4.7C upper panels). Prolonged incubation of conidia in water resulted in a concentrated punctuate localization in the cytoplasm. Interestingly, when the conidia became immobilized on the surface, a transient localization of the GFP fluorescence was observed in the nucleus (Figure 4.7C lower panels). This might to be coincident with the germination. These results suggest that the localization of MgFTT1 is dynamic and responsive to environmental stimuli.

To evaluate whether the over-expression affects infection-related development, conidia were incubated on glass cover slips and on onion peels (Figure 4.8). Appressoria were evident. However, some formed balloon like structures which were not melanized. Many of these resumed polarized growth (Figure 4.8A). Appressoria appeared to be larger than wild type, and presented more chitinous material compared to the wild type as evidenced by calcofluor staining (Figure 4.8B and 4.8C). Similar structures were also observed on onion peels. Thus, MgFTT1 appears to be involved in the initiation of germination and maintenance of filamentous growth, which is supported by both Northern analysis and intracellular localization studies in addition to being involved in nutrient sensing.

Discussion

14-3-3 proteins mediate various cellular processes. Despite the presence of the gene in all eukaryotes studied so far, its role is not clearly defined. In higher eukaryotes,

this, in part, may be due to functional redundancy. In work presented here, we characterized a 14-3-3 homolog (MgFTT1) in the rice blast fungus. The phylogenetic analysis and protein domain information confirm that MgFTT1 is a 14-3-3 protein. The phylogenetic analysis showed the presence of two putative isoforms in genomes of filamentous fungi, which may suggest isoform specific function (Rosenquist et al., 2000; Roberts and de Bruxelles, 2002; Daugherty et al., 1996). Thus, it could be anticipated that functional analysis of the other copy of the gene in the rice blast fungus might reveal different phenotypes.

The expression pattern of the transcript appears to be finely tuned. However, evidence of transcriptional regulation of 14-3-3 is limited (van Hemert et al., 2001; Hurtado and Rachubinski, 2002). In our study, transcript levels of MgFTT1 were higher in polarized growth stages such as germination and mycelial growth than in conidia and appressoria. The transcriptional variation detected in our Northern analysis is consistent with the functions of this gene we observed in *M. grisea*. Y1BMH1, a 14-3-3 gene in *Y. lipolytica*, was shown to be upregulated during yeast-to-hyphal transition and germination (Hurtado and Rachubinski, 2002). Thus, 14-3-3 appears to regulate polarized growth. Also, the transcription of MgFTT1 was regulated by starvation stress. Consistent with the role of the 14-3-3 proteins in other organisms, transcription was affected by carbon starvation. Multi-copy overexpression of yeast BMH1 in *S. cerevisiae* resulted in poor growth in a medium with a poor carbon source (van Heusden et al., 1992). Similar to observations in budding yeast, 14-3-3 regulates carbon nutrient utilization in gene copy number dependent manner in *C. albicans* (Wang et al., 2004). Thus, it is possible that

stoichiometric balance of 14-3-3 protein with interacting partners is necessary for the appropriate physiological response.

The effect on the nutrient metabolism

Deletion of MgFTT1 from the rice blast fungus resulted in increased activities for some hydrolytic enzymes that may be under transcriptional induction or catabolite repression. Higher β -galactosidase, β -glucosidase, and invertase activities were shown in the culture filtrate of deletion mutants compared to the wild type strain. In budding yeast, β -galactosidase and invertase are under catabolite repression. These results suggest that MgFTT1 is involved in catabolite repression mediated gene regulation. In budding yeast, treatment of rapamycin which mimics nutrient deprivation triggered the dissociation of MSN2/4 transcription factors from 14-3-3, and caused the transcription factors to translocate into the nucleus where they bind stress response element (SRE). The role of 14-3-3 protein in yeast species in nutrient sensitive regulation has been shown in other studies (van Heusden et al., 1992; Beck and Hall, 1999; Bertram et al., 1998; Wang et al., 2004; Ichimura et al., 2004). A large scale proteomic and expression profiling study showed that the deletion events led to the increased expression level of genes that are under carbon and nitrogen regulatory pathways (Ichimura, Kubota, et al. 2004 115 /id). The authors also noted the possible negative regulation of ste20 pathway by BMH1 and BMH2, affecting downstream processes of intermediary metabolism and cell wall integrity. A similar study also suggested that 14-3-3 may affect downstream targets by negative regulation in budding yeast (Bruckmann et al., 2004). Also, the study indicated that 14-3-3s regulate stress-sensitive gene expression, suggesting 14-3-3s' role in general

stress response. In *Cochliobolus carbonum*, deletion of the SNF kinase did not affect invertase activity suggesting it is not under catabolite repression (Tonukari et al., 2000). In our experiments, deletion of MgFTT1 resulted in higher invertase activity than the wild type. Since sucrose is one of the main sugars for transport in plants, invertase activity might not be under strict regulation of catabolite repression or different isozymes that are under catabolite repression may contribute to the total invertase activity for plant pathogens.

Appressorium formation

The deletion of MgFTT1 did not appear to affect appressorium formation on an inductive surface. Also, chemical inducers such as cAMP, IBMX, and cutin monomers were effective in inducing appressoria in the deletion mutants. The deletion mutants also formed appressoria on plant surfaces such as onion skin and rice leaves. Thus, it appears that even though the gene is involved in polarized growth, the ability to form appressoria was not hampered by the deletion of the gene. This might be ascribed to functional redundancy of another copy of 14-3-3 in the rice blast fungus genome. In the rice blast fungus, PMK1 MAP kinase is essential for the formation of the appressorium (Xu and Hamer, 1996). In budding yeast, both copies of 14-3-3 proteins were shown to bind STE20. The physical interaction of MAPKKK with 14-3-3 proteins appears to be a common mechanism in eukaryotes. Recent work to dissect PMK1 MAPK pathway suggests that there might be divergent input and output of the signals despite the conserved structural modules. The characterization of MST12, the STE12 homolog of budding yeast, in the rice blast fungus revealed the presence of possible PKA

phosphorylation site and divergence of gene structure in filamentous fungi compared to the budding yeast (Park et al., 2004). Also, deletion of the PAK kinase homologs in the rice blast fungus MST20 and CHM1, homologs of STE20 and CLA4 in budding yeast, respectively, showed phenotypes that are different from the PMK1 mutant (Li et al., 2004). Thus, it appears appressorium formation in the rice blast fungus is regulated by the cooperation of cAMP and PMK1-related MAPK pathways that is conserved in filamentous fungi. However, the role of 14-3-3 proteins in the regulation of MAPK pathway in *M. grisea* remains to be established.

Cell wall remodeling

14-3-3 proteins appear to be involved in cell wall integrity in fungi, which may be mediated by various signaling components. When MgFTT1 was deleted, we detected less staining of chitinous material by WGA staining. On the other hand, the over-expression strain showed an increased staining of chitinous material compared to the wild type. One of the important pathways for cell wall remodeling in fungi is through protein kinase C and Rac1. A recent finding in fission yeast suggests that rad25 plays a repressor of cell wall defectiveness, and multicopy expression affects cell morphology and cell wall properties (Ishiguro et al., 2001). Interestingly, a deletion mutant of Cla4, a PAK kinase homolog in *U. maydis*, showed delocalized chitin deposition in the cell wall (Leveleki et al., 2004). Also, 14-3-3 homologs in fission yeast, rad24 and rad25, were shown to interact with Byr2, a MAPKKK, and act as negative regulators by affecting the localization of the BYR2 protein (Ozoe et al., 2002). Thus, there appears to be a common mechanism of the regulation of the MAPK pathway by 14-3-3 proteins. Preliminary yeast

two hybrid analysis with MgFTT1 as bait and MST11, a STE11 homolog of budding yeast, suggested no physical interaction in the rice blast fungus (Jeong and Dean, Unpublished data). The genetic and physical interactions of 14-3-3 with STE20 PAK kinase and MAPKKKs suggest that the role of the 14-3-3 protein in the signaling cascade is to bring the components into proximity for interaction, acting as a scaffold. PAK kinases are downstream effectors of Rac and Cdc42 through actin cytoskeleton reorganization (Eby et al., 1998; Holly and Blumer, 1999). Interestingly, BMH2 and BMH1, to a lesser degree, can rescue the clathrin heavy chain mediated vesicular transport in yeast (Gelperin et al., 1995). These results suggest that 14-3-3 protein is a critical regulator of the actin cytoskeleton. The different routes by which the 14-3-3 proteins were identified further support this conserved role. As mentioned previously, YIBMH1 was cloned as a complementation to the Rac1 deficient phenotype in *Y. lipolytica* (Hurtado and Rachubinski, 2002). Also, Cla4 was suggested to be the downstream target of Rac1 in *U. maydis* (Leveleki et al., 2004). Rac1 is known to interact with CRIB domain in PAK kinases. Although the physical interaction of 14-3-3 proteins and STE20 PAK kinase was shown in yeast, genetic interaction of 14-3-3 proteins and Rac1/cdc42 appears to be indirect (Zenke et al., 2004).

Pathogenicity

Disease assays showed that the deletion of MgFTT1 resulted in reduced pathogenicity. Occasional lesions were observed on inoculated rice plants. However, lesions were localized, and progressed more slowly compared to the lesions caused by the wild type. Deletion mutants were able to penetrate an artificial surface. The ability to

penetrate cellophane film appears to represent the ability to form functional appressorium since *mac1* (no appressorium formation), *cpka* (delayed and non-functional appressorium formation), and *pmk1* (no appressorium formation and no invasive growth) were not able to penetrate and grow. This was further corroborated by the wounded barley leaf assay. In contrast to the wild type, the deletion mutant strains showed limited symptom development in wounded barley leaves, suggesting a defect in post-penetration growth. Thus, it appears that the deletion of the gene affects *in planta* growth. As described previously, deletion of *MgFTT1* resulted in the alteration of cell wall property. Chitin is an important component of fungal cell wall. In fission yeast, one of 14-3-3 homolog, *rad25*, was isolated as by complementing a mutant strain with defects in cell wall integrity (Ishiguro et al., 2001). Interestingly, overexpression of *rad25* resulted in more intense staining of septum with calcofluor, also the overexpression strain exhibited abnormal elongation of cell morphology. Thus, it is reasonable to assume that deletion of *MgFTT1* results in a cell wall defect that affects *in planta* growth. Also, *MgFTT1* deletion strains exhibited alteration in the activity of some carbohydrate hydrolyzing enzymes. Some of these enzymes or enzymatic product might represent pathogen-associated molecular pattern (PAMP). Chitin metabolism appears to be elevated during appressorium formation compared to mycelium grown in rich medium and carbon starvation condition (unpublished data). Chitin deacetylases are up-regulated during carbon starvation and toward late stages of appressoria on an artificial surface which might represent the penetration stage. In fact, carbon utilization pathways may be primary targets of *MgFTT1* as shown in this work. Thus, *MgFTT1* deletion strains might present more PAMPs, which can be recognized by host defense systems. Deletion of *MgFTT1*

resulted in less intense staining with WGA compared to wild type, suggesting reduced level of chitin in the cell wall. Chitin binding domain proteins have been proposed to protect the fungal cell wall from chitinases produced by plant host. Also, chitin derivatives such as chitosan have been proposed to be an effective way to protect the fungal cell wall from host chitinases during penetration (Mendgen et al., 1996). Thus, alteration in chitin metabolism may not only weaken the cell wall, but may also make the fungus more susceptible to host defense responses, resulting in greatly reduced ability to grow *in planta*.

Localization of MgFTT1 GFP fusion

To examine the temporal and spatial localization of 14-3-3, we created transformants in which MgFTT1 was tagged at the C-terminus with GFP. In many organisms, 14-3-3 localizes to the nucleus. In a large scale protein localization analysis in yeast, BMH2 was found to localize to the nucleus under certain conditions (Huh et al., 2003). Also, in tobacco, a 14-3-3 protein localizes to the nucleus (Igarashi et al., 2001). Our finding that MgFTT1 is excluded from the nucleus in conidia and exhibit transient nuclear localization, presumably, before germination may suggest that it shuttles repressor activity out of the nucleus. This appears to be the case in tobacco. A bZIP transcription factor RSG is negatively regulated by a 14-3-3 protein via its intracellular localization. Also, one of the proposed roles of 14-3-3 protein is in nuclear export by masking nuclear localizing signals (NLS), thus reducing the rate of nuclear import of target proteins (MacKintosh, 2004). In *Arabidopsis*, 14-3-3 proteins show isotype-specific and tissue-specific localization. Some isotypes show nuclear localization (Paul et

al., 2005; Bihn et al., 1997). However, a detailed localization study in *Arabidopsis* suggests that the subcellular localization of 14-3-3 proteins is modulated by interaction partners rather than by 14-3-3 protein itself (Paul et al., 2005). Detailed studies support a complementary role of 14-3-3 in nucleocytoplasmic shuttling of interacting proteins through physical binding (Brunet et al., 2002). Thus, the biological significance of the nuclear cytoplasmic shuttling remains to be determined.

Interaction with other signal pathways

In many organisms, 14-3-3 appears to regulate polarized growth and reinitiation of growth or germination in a variety of contexts (Morton et al., 2002; Roberts and Mosch, 1997; Igarashi et al., 2001; Kraus et al., 2002; Benton et al., 2002). A variety of evidence indicate that 14-3-3 is involved in cross-talk between signal transduction pathways (Dumaz and Marais, 2005). Physical interactions of 14-3-3 with MAPK component such as MAPKKK, raf, which is an established MAPKKK, and its upstream component, a PAK, Ste20 have been shown (Irie et al., 1994; Roberts and Mosch, 1997; Ozoe et al., 2002). 14-3-3 is known to be located as a down stream regulator of the TOR pathway (Bertram et al., 1998). TOR is a key regulator of cell growth, and interacts with other signaling pathway such as cAMP/PKA. In this work, we noted alterations of the cell wall properties of the mutants compared to wild type. Also, an aberrant morphology was seen in the overexpression strain. Recent microarray study suggested that 14-3-3s play regulatory roles in carbon and nitrogen responsive cellular responses through PKA, SNF1, and TOR pathways in yeast (Ichimura et al., 2004). Proteomic studies also suggested roles of 14-3-3 proteins in cytoskeleton dynamics, intracellular trafficking, and

signaling pathways (Jin et al., 2004; Pozuelo et al., 2004). Considering the cellular processes that 14-3-3s are involved in, phenotypes of knockout strains of individual 14-3-3 genes in yeast and the rice blast fungus appear to be relatively subtle. This suggests that there might be functional overlap between 14-3-3s in particular organisms. However, various lines of evidence suggest independent roles exist for individual 14-3-3s.

In this work, we characterized MgFTT1, a 14-3-3 homolog in the rice blast fungus. Our results showed that 14-3-3 protein is an important component of cellular morphogenesis and physiology. MgFTT1 was shown to play a critical role in establishing pathogenic interaction with host plant. Also, MgFTT1 affects cellular physiology which might affect the overall infection process. Considering the possible existence of two isoforms in filamentous fungi, functional analysis of the other isoform awaits further investigation.

References

1. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W., and Lipman,D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402.
2. Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A., and Struhl,K. (1987). *Current protocols in molecular biology*. John Wiley & Sons, Inc.).
3. Banuett,F. (1998). Signalling in the yeasts: An informational cascade with links to the filamentous fungi. *Microbiol. Mol. Biol. Rev.* 62, 249-274.
4. Beck,T. and Hall,M.N. (1999). The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors. *Nature* 402, 689-692.
5. Benton,R., Palacios,I.M., and St Johnston,D. (2002). *Drosophila* 14-3-3/PAR-5 is an essential mediator of PAR-1 function in axis formation. *Dev. Cell* 3, 659-671.
6. Bertram,P.G., Zeng,C., Thorson,J., Shaw,A.S., and Zheng,X.F. (1998). The 14-3-3 proteins positively regulate rapamycin-sensitive signaling. *Curr. Biol* 8, 1259-1267.

7. Bihn,E.A., Paul,A.L., Wang,S.W., Erdos,G.W., and Ferl,R.J. (1997). Localization of 14-3-3 proteins in the nuclei of arabidopsis and maize. *Plant J* 12, 1439-1445.
8. Both,M., Csukai,M., Stumpf,M.P., and Spanu,P.D. (2005). Gene expression profiles of *Blumeria graminis* indicate dynamic changes to primary metabolism during development of an obligate biotrophic pathogen. *Plant Cell* 17, 2107-2122.
9. Bourett,T.M., Sweigard,J.A., Czymmek,K.J., Carroll,A., and Howard,R.J. (2002). Reef coral fluorescent proteins for visualizing fungal pathogens. *Fungal Genet. Biol* 37, 211-220.
10. Bruckmann,A., Steensma,H.Y., Teixeira De Mattos,M.J., and van Heusden,G.P. (2004). Regulation of transcription by *Saccharomyces cerevisiae* 14-3-3 proteins. *Biochem J* 382, 867-875.
11. Brunet,A., Kanai,F., Stehn,J., Xu,J., Sarbassova,D., Frangioni,J.V., Dalal,S.N., DeCaprio,J.A., Greenberg,M.E., and Yaffe,M.B. (2002). 14-3-3 transits to the nucleus and participates in dynamic nucleocytoplasmic transport. *J Cell Biol* 156, 817-828.
12. Carroll,A.M., Sweigard,J.A., and Valent,B. (1994). Improved vectors for selecting resistance to hygromycin. *Fungal Genet Newslett* 41, 22.

13. Chenna,R., Sugawara,H., Koike,T., Lopez,R., Gibson,T.J., Higgins,D.G., and Thompson,J.D. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* *31*, 3497-3500.
14. Choi,W. and Dean,R.A. (1997). The adenylate cyclase gene MAC1 of *Magnaporthe grisea* controls appressorium formation and other aspects of growth and development. *Plant Cell* *9*, 1973-1983.
15. Cognetti,D., Davis,D., and Sturtevant,J. (2002). The *Candida albicans* 14-3-3 gene, BMH1, is essential for growth. *Yeast* *19*, 55-67.
16. Daugherty,C.J., Rooney,M.F., Miller,P.W., and Ferl,R.J. (1996). Molecular organization and tissue-specific expression of an *Arabidopsis* 14-3-3 gene. *Plant Cell* *8*, 1239-1248.
17. de Jong,J.C., McCormack,B.J., Smirnov,N., and Talbot,N.J. (1997). Glycerol generates turgor in rice blast. *Nature* *389*, 244-245.
18. Dean,R.A., Talbot,N.J., Ebbole,D.J., Farman,M.L., Mitchell,T.K., Orbach,M.J., Thon,M., Kulkarni,R., Xu,J.R., Pan,H., Read,N.D., Lee,Y.H., Carbone,I., Brown,D., Oh,Y.Y., Donofrio,N., Jeong,J.S., Soanes,D.M., Djonovic,S., Kolomiets,E., Rehmeyer,C., Li,W., Harding,M., Kim,S., Lebrun,M.H., Bohnert,H., Coughlan,S., Butler,J., Calvo,S., Ma,L.J., Nicol,R., Purcell,S.,

- Nusbaum,C., Galagan,J.E., and Birren,B.W. (2005). The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* 434, 980-986.
19. DeZwaan,T.M., Carroll,A.M., Valent,B., and Sweigard,J.A. (1999). *Magnaporthe grisea* Pth11p is a novel plasma membrane protein that mediates appressorium differentiation in response to inductive surface cues. *Plant Cell* 11, 2012-2030.
 20. Dixon,K.P., Xu,J.R., Smirnov,N., and Talbot,N.J. (1999). Independent signaling pathways regulate cellular turgor during hyperosmotic stress and appressorium-mediated plant infection by *Magnaporthe grisea*. *Plant Cell* 11, 2045-2058.
 21. Dumaz,N. and Marais,R. (2005). Integrating signals between cAMP and the RAS/RAF/MEK/ERK signalling pathways. *FEBS J* 272, 3491-3504.
 22. Eby,J.J., Holly,S.P., van Drogen,F., Grishin,A.V., Peter,M., Drubin,D.G., and Blumer,K.J. (1998). Actin cytoskeleton organization regulated by the PAK family of protein kinases. *Curr. Biol* 8 , 967-970.
 23. Fu,H., Subramanian,R.R., and Masters,S.C. (2000). 14-3-3 proteins: structure, function, and regulation. *Annu. Rev Pharmacol. Toxicol.* 40, 617-647.
 24. Galagan,J.E., Calvo,S.E., Borkovich,K.A., Selker,E.U., Read,N.D., Jaffe,D., FitzHugh,W., Ma,L.J., Smirnov,S., Purcell,S., Rehman,B., Elkins,T., Engels,R., Wang,S., Nielsen,C.B., Butler,J., Endrizzi,M., Qui,D., Ianakiev,P., Bell-

- Pedersen,D., Nelson,M.A., Werner-Washburne,M., Selitrennikoff,C.P.,
Kinsey,J.A., Braun,E.L., Zelter,A., Schulte,U., Kothe,G.O., Jedd,G., Mewes,W.,
Staben,C., Marcotte,E., Greenberg,D., Roy,A., Foley,K., Naylor,J., Stange-
Thomann,N., Barrett,R., Gnerre,S., Kamal,M., Kamvysselis,M., Mauceli,E.,
Bielke,C., Rudd,S., Frishman,D., Krystofova,S., Rasmussen,C., Metzenberg,R.L.,
Perkins,D.D., Kroken,S., Cogoni,C., Macino,G., Catcheside,D., Li,W., Pratt,R.J.,
Osmani,S.A., DeSouza,C.P.C., Glass,L., Orbach,M.J., Berglund,J.A., Voelker,R.,
Yarden,O., Plamann,M., Seiler,S., Dunlap,J., Radford,A., Aramayo,R.,
Natvig,D.O., Alex,L.A., Mannhaupt,G., Ebbole,D.J., Freitag,M., Paulsen,I.,
Sachs,M.S., Lander,E.S., Nusbaum,C., and Birren,B. (2003). The genome
sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422, 859-868.
25. Gancedo,J.M. (1998). Yeast carbon catabolite repression. *Microbiol Mol Biol Rev*
62, 334-361.
26. Garrington,T.P. and Johnson,G.L. (1999). Organization and regulation of
mitogen-activated protein kinase signaling pathways. *Curr. Opin. Cell Biol* 11,
211-218.
27. Gelperin,D., Weigle,J., Nelson,K., Roseboom,P., Irie,K., Matsumoto,K., and
Lemmon,S. (1995). 14-3-3 proteins: potential roles in vesicular transport and Ras
signaling in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A* 92, 11539-
11543.

28. Gilbert,R.D., Johnson,A.M., and Dean,R.A. (1996). Chemical signals responsible for appressorium formation in the rice blast fungus *Magnaporthe grisea*. *Physiol Mol Plant P* 48, 335-346.
29. Holly,S.P. and Blumer,K.J. (1999). PAK-family kinases regulate cell and actin polarization throughout the cell cycle of *Saccharomyces cerevisiae*. *J. Cell Biol.* 147, 845-856.
30. Huh,W.K., Falvo,J.V., Gerke,L.C., Carroll,A.S., Howson,R.W., Weissman,J.S., and O'Shea,E.K. (2003). Global analysis of protein localization in budding yeast. *Nature* 425, 686-691.
31. Hurtado,C.A.R. and Rachubinski,R.A. (2002). Y1BMH1 encodes a 14-3-3 protein that promotes filamentous growth in the dimorphic yeast *Yarrowia lipolytica*. *Microbiology* 148, 3725-3735.
32. Ichimura,T., Kubota,H., Goma,T., Mizushima,N., Ohsumi,Y., Iwago,M., Kakiuchi,K., Shekhar,H.U., Shinkawa,T., Taoka,M., Ito,T., and Isobe,T. (2004). Transcriptomic and proteomic analysis of a 14-3-3 gene-deficient yeast. *Biochemistry* 43, 6149-6158.
33. Igarashi,D., Ishida,S., Fukazawa,J., and Takahashi,Y. (2001). 14-3-3 proteins regulate intracellular localization of the bZIP transcriptional activator RSG. *Plant Cell* 13, 2483-2497.

34. Irie,K., Gotoh,Y., Yashar,B.M., Errede,B., Nishida,E., and Matsumoto,K. (1994). Stimulatory effects of yeast and mammalian 14-3-3 proteins on the Raf protein kinase. *Science* 265, 1716-1719.
35. Ishiguro,J., Shimada,S., Gabriel,M., and Kopecka,M. (2001). Characterization of a fission yeast mutant which displays defects in cell wall integrity and cytokinesis. *Genes Genet Syst.* 76, 257-269.
36. Jin,J., Smith,F.D., Stark,C., Wells,C.D., Fawcett,J.P., Kulkarni,S., Metalnikov,P., O'Donnell,P., Taylor,P., Taylor,L., Zougman,A., Woodgett,J.R., Langeberg,L.K., Scott,J.D., and Pawson,T. (2004). Proteomic, functional, and domain-based analysis of *in vivo* 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization. *Curr. Biol* 14, 1436-1450.
37. Kraus,P.R., Hofmann,A.F., and Harris,S.D. (2002). Characterization of the *Aspergillus nidulans* 14-3-3 homologue, ArtA. *FEMS Microbiol Lett.* 210, 61-66.
38. Kumar,S., Tamura,K., and Nei,M. (2004). MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief. Bioinform.* 5, 150-163.
39. Lee,J., Kwon,K.S., and Hah,Y.C. (1996). Regulation of beta-glucosidase biosynthesis in *Aspergillus nidulans*. *FEMS Microbiol Lett.* 135, 79-84.

40. Lee, Y.H. and Dean, R.A. (1993). cAMP regulates infection structure formation in the plant pathogenic fungus *Magnaporthe grisea*. *Plant Cell* 5, 693-700.
41. Leung, H., Lehtinen, U., Karjalainen, R., Skinner, D., Tooley, P., Leong, S., and Ellingboe, A. (1990). Transformation of the rice blast fungus *Magnaporthe grisea* to hygromycin B resistance. *Curr. Genet.* 17, 409-411.
42. Leveleki, L., Mahlert, M., Sandrock, B., and Bolker, M. (2004). The PAK family kinase Cla4 is required for budding and morphogenesis in *Ustilago maydis*. *Mol Microbiol* 54, 396-406.
43. Li, L., Xue, C., Bruno, K., Nishimura, M., and Xu, J.R. (2004). Two PAK kinase genes, CHM1 and MST20, have distinct functions in *Magnaporthe grisea*. *Mol Plant Microbe Interact.* 17, 547-556.
44. Lopez-Girona, A., Furnari, B., Mondesert, O., and Russell, P. (1999). Nuclear localization of Cdc25 is regulated by DNA damage and a 14-3-3 protein. *Nature* 397, 172-175.
45. MacKintosh, C. (2004). Dynamic interactions between 14-3-3 proteins and phosphoproteins regulate diverse cellular processes. *Biochem J* 381, 329-342.
46. Mendgen, K., Hahn, M., and Deising, H. (1996). Morphogenesis and mechanisms of penetration by plant pathogenic fungi. *Annu. Rev Phytopathol.* 34, 367-386.

47. Mitchell,T.K. and Dean,R.A. (1995). The cAMP-dependent protein kinase catalytic subunit is required for appressorium formation and pathogenesis by the rice blast pathogen *Magnaporthe grisea*. *Plant Cell* 7, 1869-1878.
48. Morton,D.G., Shakes,D.C., Nugent,S., Dichoso,D., Wang,W., Golden,A., and Kempfues,K.J. (2002). The *Caenorhabditis elegans* par-5 gene encodes a 14-3-3 protein required for cellular asymmetry in the early embryo. *Dev. Biol* 241, 47-58.
49. Nakai,K. and Horton,P. (1999). PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci* 24, 34-36.
50. Ozoe,F., Kurokawa,R., Kobayashi,Y., Jeong,H.T., Tanaka,K., Sen,K., Nakagawa,T., Matsuda,H., and Kawamukai,M. (2002). The 14-3-3 proteins Rad24 and Rad25 negatively regulate Byr2 by affecting its localization in *Schizosaccharomyces pombe*. *Mol. Cell. Biol.* 22, 7105-7119.
51. Park,G., Bruno,K.S., Staiger,C.J., Talbot,N.J., and Xu,J.R. (2004). Independent genetic mechanisms mediate turgor generation and penetration peg formation during plant infection in the rice blast fungus. *Mol Microbiol* 53, 1695-1707.
52. Paul,A.L., Sehnke,P.C., and Ferl,R.J. (2005). Isoform-specific subcellular localization among 14-3-3 proteins in *Arabidopsis* seems to be driven by client interactions. *Mol Biol Cell* 16, 1735-1743.

53. Pozuelo,R.M., Geraghty,K.M., Wong,B.H., Wood,N.T., Campbell,D.G., Morrice,N., and MacKintosh,C. (2004). 14-3-3-affinity purification of over 200 human phosphoproteins reveals new links to regulation of cellular metabolism, proliferation and trafficking. *Biochem J* 379, 395-408.
54. Roberts,M.R. and de Bruxelles,G.L. (2002). Plant 14-3-3 protein families: evidence for isoform-specific functions? *Biochem Soc. Trans.* 30, 373-378.
55. Roberts,R.L. and Mosch,H.U. (1997). 14-3-3 proteins are essential for RAS/MAPK cascade signaling during pseudohyphal development in *S. cerevisiae*. *Cell* 89, 1055-1065.
56. Rosenquist,M., Sehnke,P., Ferl,R.J., Sommarin,M., and Larsson,C. (2000). Evolution of the 14-3-3 protein family: does the large number of isoforms in multicellular organisms reflect functional specificity? *J. Mol. Evol.* 51, 446-458.
57. Roth,D., Birkenfeld,J., and Betz,H. (1999). Dominant-negative alleles of 14-3-3 proteins cause defects in actin organization and vesicle targeting in the yeast *Saccharomyces cerevisiae*. *FEBS Lett.* 460, 411-416.
58. Szabo,R. (2001). Cla4 protein kinase is essential for filament formation and invasive growth of *Yarrowia lipolytica*. *Mol Genet Genomics* 265, 172-179.

59. Tonukari,N.J., Scott-Craig,J.S., and Walton,J.D. (2000). The *Cochliobolus carbonum* SNF1 gene is required for cell wall-degrading enzyme expression and virulence on maize. *Plant Cell* *12*, 237-248.
60. van Hemert,M.J., van Heusden,G.P., and Steensma,H.Y. (2001). Yeast 14-3-3 proteins. *Yeast* *18*, 889-895.
61. van Heusden,G.P., Wenzel,T.J., Lagendijk,E.L., de Steensma,H.Y., and van den Berg,J.A. (1992). Characterization of the yeast BMH1 gene encoding a putative protein homologous to mammalian protein kinase II activators and protein kinase C inhibitors. *FEBS Lett.* *302*, 145-150.
62. Vasara,T., Keranen,S., Penttila,M., and Saloheimo,M. (2002). Characterisation of two 14-3-3 genes from *Trichoderma reesei*: interactions with yeast secretory pathway components. *Biochim Biophys Acta.* *1590*, 27-40.
63. Wang,Y.K., Das,B., Huber,D.H., Wellington,M., Kabir,M.A., Sherman,F., and Rustchenko,E. (2004). Role of the 14-3-3 protein in carbon metabolism of the pathogenic yeast *Candida albicans*. *Yeast* *21*, 685-702.
64. Xu,J.R. (2000). MAP kinases in fungal pathogens. *Fungal Genet. Biol* *31*, 137-152.

65. Xu,J.R. and Hamer,J.E. (1996). MAP kinase and cAMP signalling regulate infection structure formation and pathogenic growth in the rice blast fungus *Magnaporthe grisea*. *Genes Dev.* *10*, 2696-2706.
66. Xu,J.R., Urban,M., Sweigard,J.A., and Hamer,J.E. (1997). The CPKA gene of *Magnaporthe grisea* is essential for appressorial penetration. *Mol Plant Microbe Interact.* *10*, 187-194.
67. Yaffe,M.B. (2002). How do 14-3-3 proteins work? - Gatekeeper phosphorylation and the molecular anvil hypothesis. *FEBS Lett.* *513*, 53-57.
68. Yaffe,M.B. and Elia,A.E.H. (2001). Phosphoserine/threonine-binding domains. *Curr Opin Cell Biol.* *13*, 131-138.
69. Zenke,F.T., Krendel,M., DerMardirossian,C., King,C.C., Bohl,B.P., and Bokoch,G.M. (2004). p21-activated kinase 1 phosphorylates and regulates 14-3-3 binding to GEF-H1, a microtubule-localized Rho exchange factor. *J Biol Chem.* *279*, 18392-18400.

Figure legends

Figure 4.1 Phylogenetic tree using fungal 14-3-3 RefSeq sequences retrieved from GenBank.

Maximum parsimony tree was constructed from alignment generated using CLUSTALW integrated in MEGA3 software, and the consensus tree is presented. Bootstrap values are indicated on each node (1000 iterations). Separation of yeasts, Ascomycetes, and Basidiomycetes were noted. The tree suggests that two different isoforms may exist in ascomycetous filamentous fungi (An: *A. nidulans*, Af: *A. fumigatus*, Cn: *Cryptococcus neoformans*, Gz: *G. zeae*, MgFTT1/2: *M. grisea*, Nc: *N. crassa*, Hj: *H. jeroni*, Sc: *Saccharomyces cerevisiae*, Um: *Ustilagos maydis*).

Figure 4.2 Expression pattern of MgFTT1 in different growth stages and conditions.

The expression pattern of MgFTT1 was investigated using Northern hybridization. Total RNA was extracted from conidia (lane 1), germlings (lane 2), appressoria (lane 3), complete medium (CM) (lane 4), minimal medium (MM) (lane 5), MM-Carbon (lane 6), and MM-Nitrogen (lane 7). The expression pattern suggests that the transcription is sensitive to polarized growth and nutrients. Panel (A): Northern blot probed with MgFTT1 ORF fragment, Panel (B) ribosomal banding.

Figure 4.3 Targeted deletion of MgFTT1.

(A) Deletion strategy. Deletion plasmid pFTT3 was constructed in pCB1004 by ligation of left flanking sequence (Lf: ~1.7 kb) and right flanking sequence (Rf: ~0.5kb) from

MgFTT1. (B) Confirmation of deletion. Southern blot analysis was done using MgFTT1 ORF as probe (upper panel). Genomic DNA from each strain was digested with KpnI (lane 1: wt 70-15, lane 2: Δ MgFTT1-1, lane 3: Δ MgFTT1-2, lane 4: Δ MgFTT1-3, lane 5: ectopic 1, and lane 6: ectopic 2). In the deletion mutants, no signal was seen. In wild type and ectopic mutants, signal was present. Hybridization with hygromycin probe shows signals in all transformants except for wild type (lower panel).

Figure 4.4 Enzymatic activities in culture filtrates.

Each strain was grown in MM liquid culture with sucrose as a carbon source. Culture filtrate was taken after 24 hrs for the measurement of the enzymatic activities. Wt and ect represent wild type and ectopic transformant, respectively. (A) Invertase activity, (C) β -galactosidase activity, (C) β -glucosidase activity

Figure 4.5 Deletion of MgFTT1 affects cell wall properties.

(A) Cell wall chitin staining with WGA-TRITC conjugate. In comparison to wild type and ectopic integration strains, the deletion mutant strains showed reduced staining with WGA, suggesting lower amount of chitin in the cell wall. Similar staining properties were shown in the other deletion mutants (B) Penetration of cellophane film. In the wild type and deletion strain, growth was seen, indicating the successful penetration of the film. No growth was observed for CPKA, MAC1 and PMK1 mutants, indicating defective non-functional appressorium formation and/or penetration.

Figure 4.6 Pathogenicity assays of wild type and mutant strains.

Two milliliter of 1×10^5 /ml conidial suspension was inoculated by spraying on two weeks old rice (A) and barley (B) seedlings. The picture was taken 2 weeks post inoculation. Wild type produced well-developed spreading symptoms on rice and barley plants. However, the deletion mutants (Δ MgFTT1-1, Δ MgFTT1-2, and Δ MgFTT1-3) showed greatly reduced symptoms with lesions showing limited development compared to the wild type. A similar pattern of the disease progression was seen in barley. (C) Barley plants were drop inoculated after wounding with sterile pipette tip. In contrast to wild type and ectopic transformants which produced rapid progression of symptom development, deletion strains showed a reduced symptom development, suggesting impaired *in planta* growth.

Figure 4.7 The localization of MgFTT1::GFP fusion protein.

The subcellular localization of the fusion protein was determined using a fluorescence microscope. (A) The localization of nucleus and septum was determined using calcofluor and DAPI staining. (B) The localization of the GFP fusion protein was seen throughout the cytoplasm. The arrows indicate nuclei. The GFP signal coincides with nucleus, suggesting the nuclear localization of the protein. (C) Localization of fluorescence in the conidium. The arrow in the upper left panel shows the absence of GFP signal, and the same conidium observed by DAPI stain appear to colocalize to region devoid of GFP. In contrast, the arrow in the lower panel indicates localized concentration of GFP which colocalizes to the nucleus (left panel: FITC channel for GFP signal, right panel: DAPI channel for the nucleus).

Figure 4.8 Over-expression of MgFTT1::GFP affects appressorial development.

(A) Appressorium formation by the MgFTT1::GFP strain. Most germinating conidia formed appressoria. However, some formed a balloon like structures (indicated by arrows) which were not melanized, and re-initiated polarized growth. (B) Appressorium produced by wild type on onion epidermal peels. Fully melanized appressoria were observed. (C) Appressorial produced by MgFTT1::GFP strain on onion epidermal peels. Appressoria appeared to be fully melanized. However, the appressorium cell wall appeared to stain more heavily for chitin compared to the wild type (appressoria are indicated by arrows in (B) and (C)). (Left panels: bright field, right panels: (A) FITC channel, (B) and (C) DAPI channel).

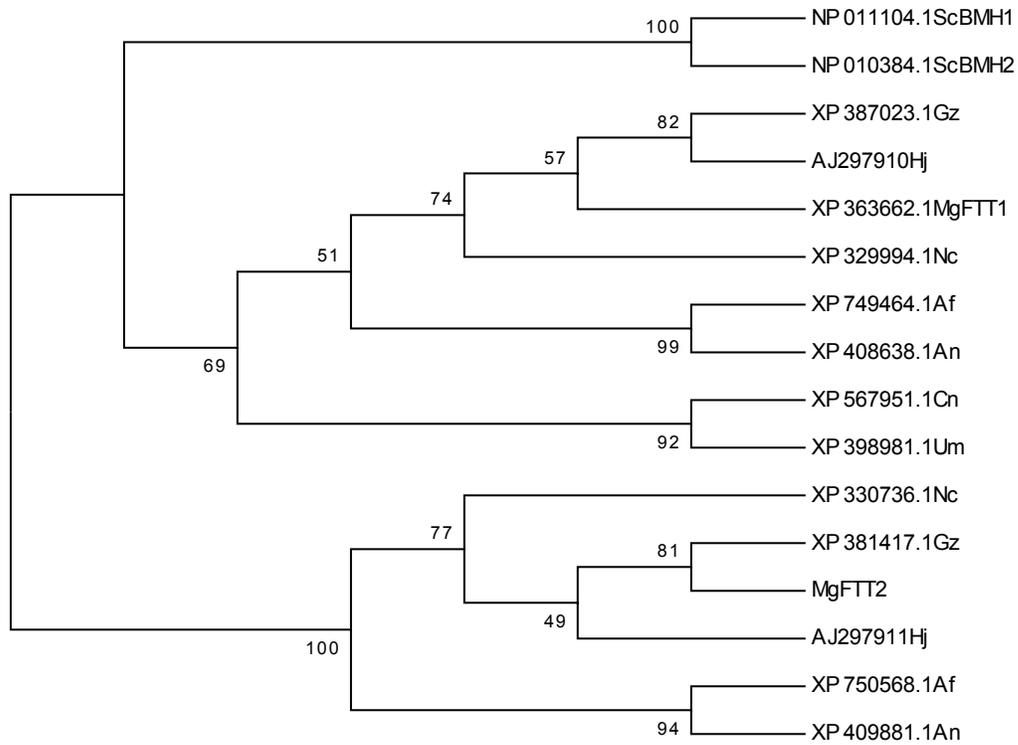


Figure 4.1 Phylogenetic tree using fungal 14-3-3 RefSeq sequences retrieved from GenBank.

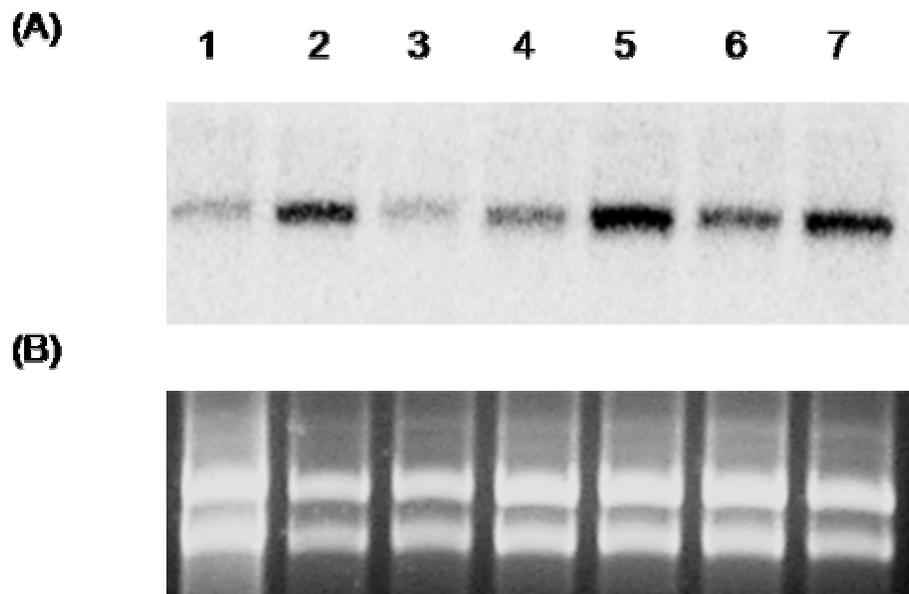


Figure 4.2 Expression pattern of MgFTT1 in different growth stages and conditions.

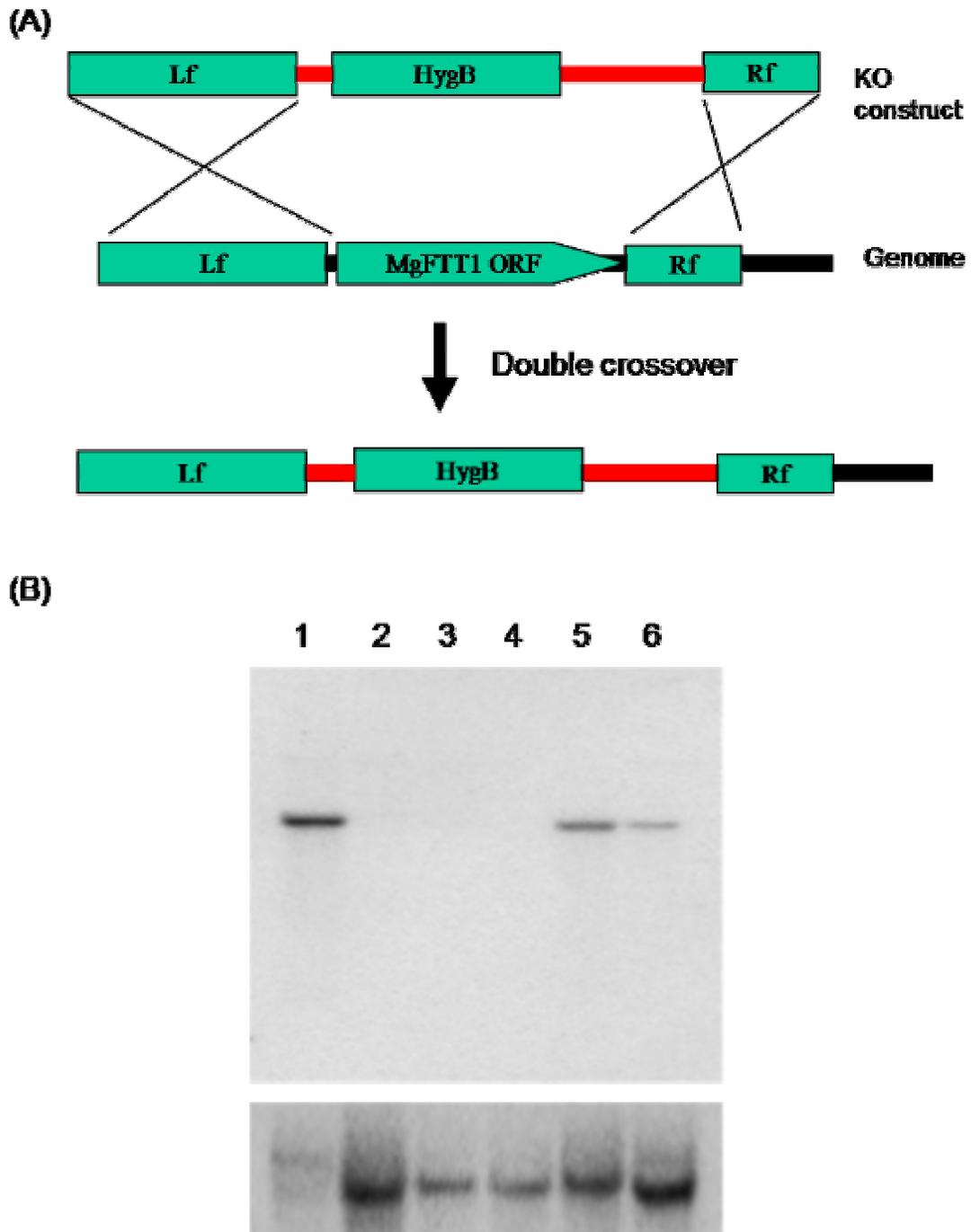
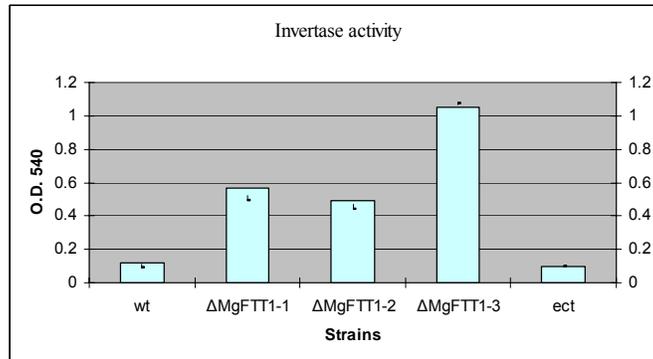
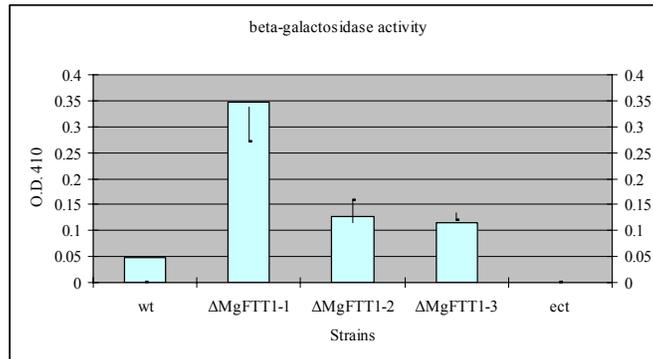


Figure 4.3 Targeted deletion of MgFTT1.

(A) Invertase activity



(B) β -galactosidase activity



(C) β -glucosidase activity

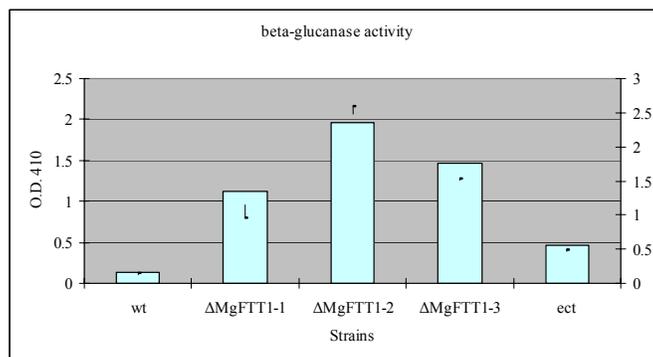


Figure 4.4 Enzymatic activities in culture filtrates.

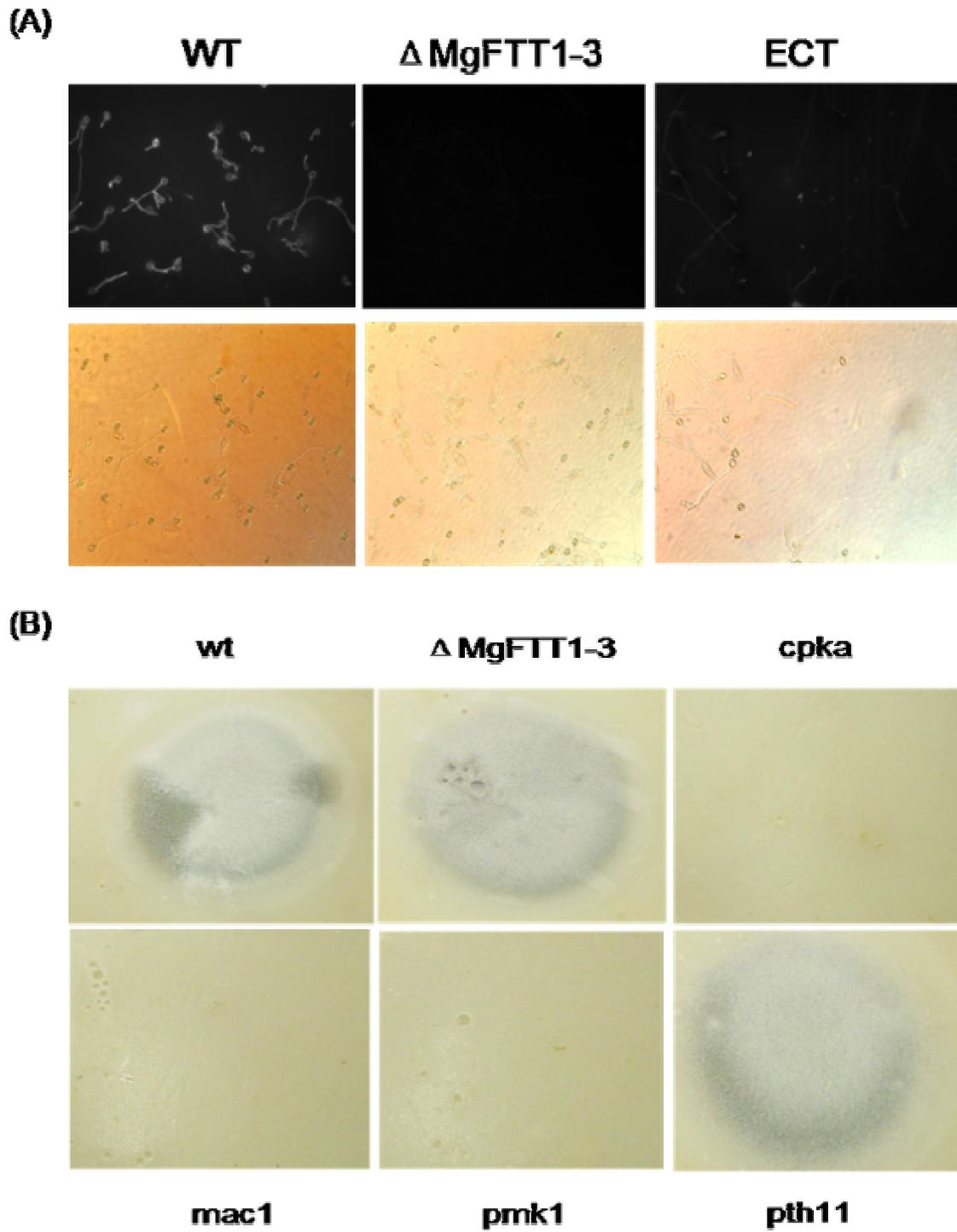


Figure 4.5 Deletion of MgFTT1 affects cell wall properties.

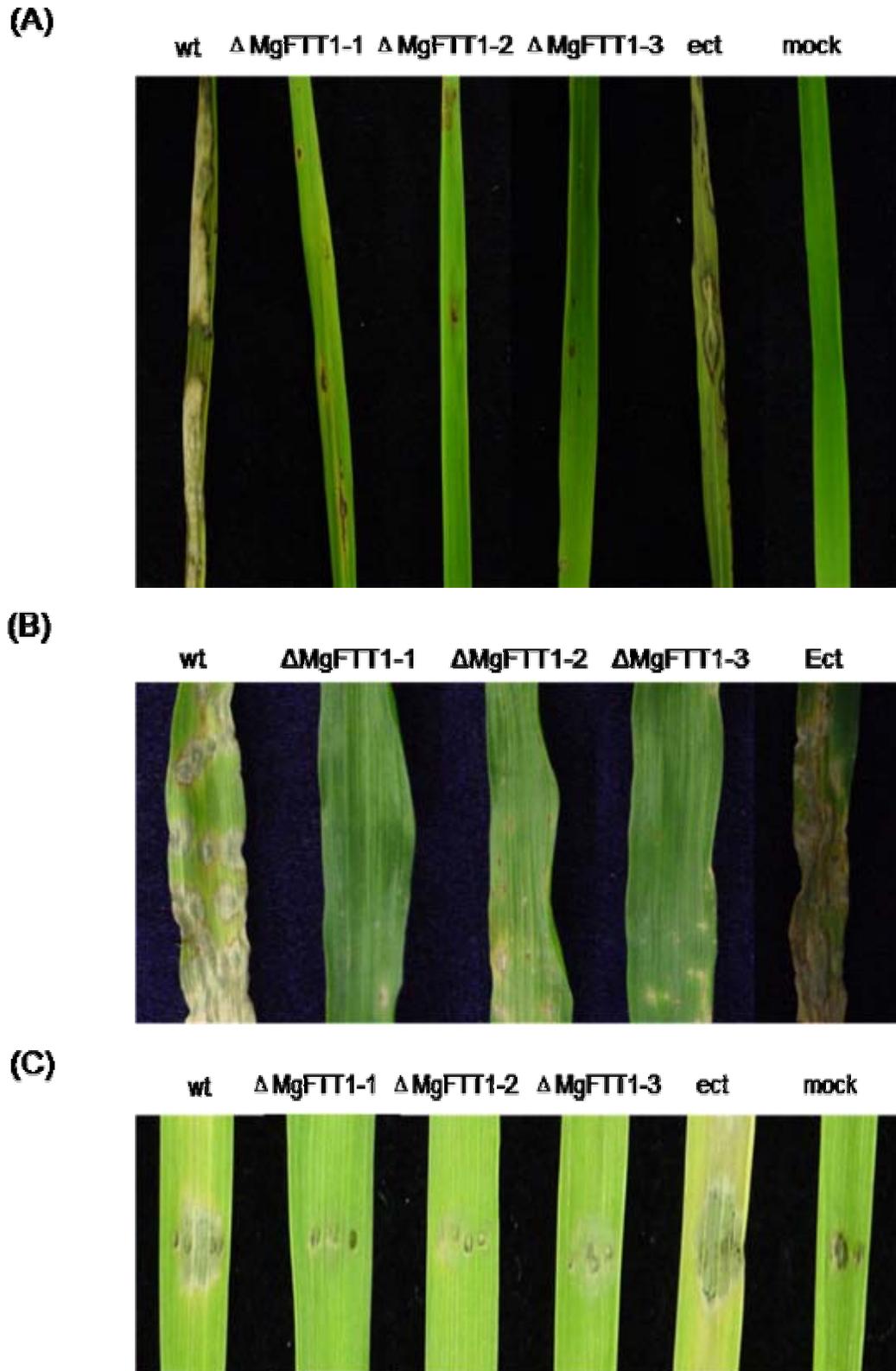


Figure 4.6 Pathogenicity assays of wild type and mutant strains.

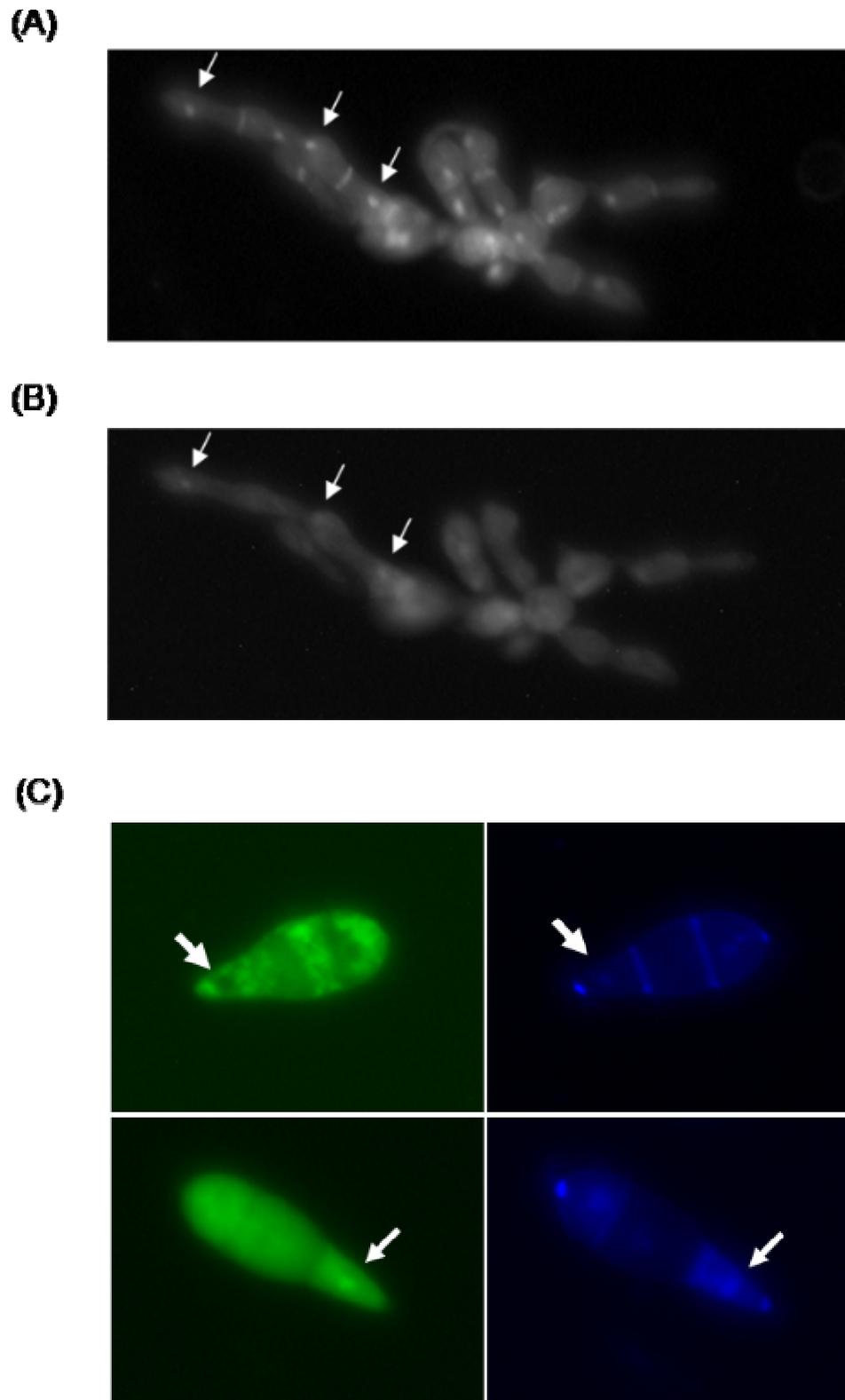
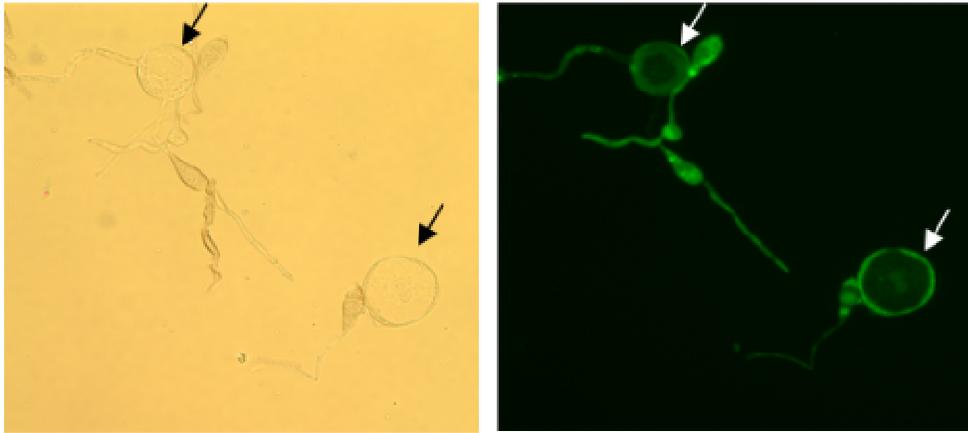
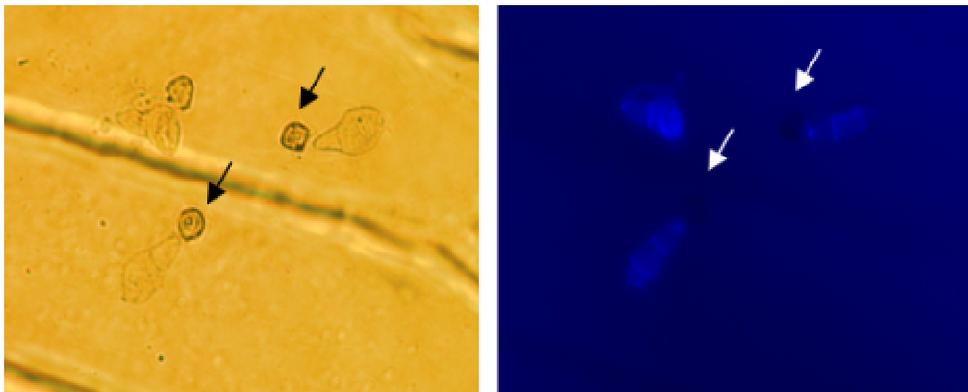


Figure 4.7 The localization of MgFTT1::GFP fusion protein.

(A)



(B)



(C)

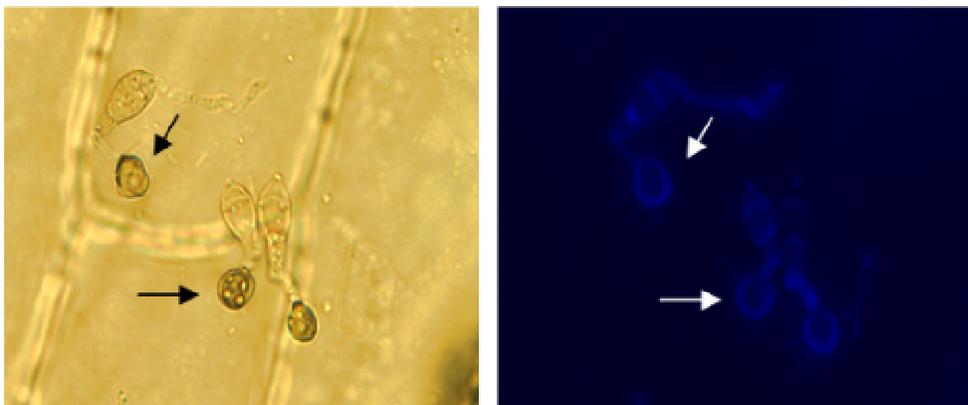


Figure 4.8 Over-expression of MgFTT1::GFP affects appressorial development.

Chapter Five

Gene disruption using split-marker in the rice blast fungus,

Magnaporthe grisea

Jun Seop Jeong and Ralph A. Dean

Center for Integrated Fungal Research (CIFR), North Carolina State University

Raleigh, NC 27695

Abstract

A number of strategies have been devised to inactivate genes in filamentous fungi. Gene knockout is a key tool for functional analysis of genes. Typically, gene deletion involves double homologous recombination replacing the gene of interest with a selectable marker. Gene deletion experiments involve creation of knockout vector with sequences flanking the gene of interest and the screening of a large number of transformants for homologous recombination. These steps can be a major bottle neck in functional gene analysis. Recently, PCR-mediated cloning-free strategies have been developed for creating constructs for gene deletion in fungi. In this work, we exploited a split-marker system for gene replacement of the *snodprot1* homolog in the rice blast fungus, *Magnaporthe grisea*. The protein is present in many filamentous fungi including *Neurospora crassa* and *Fusarium* spp.. We show that the use of split-marker approach was more efficient for the generation of deletion mutants compared to the use of traditional gene knockout constructs. Significantly, the deletion mutants showed greatly reduced pathogenicity on host plants. The reduced pathogenicity might be due to impaired growth *in planta*. *Snodprot1* homologs have been shown to be phytotoxic in other systems, and suggested to have characteristics of hydrophobins. However, our data suggested that the purified protein did not show apparent phytotoxicity when applied to wounded barley leaves. This is the first report, to our knowledge, directly implicating a *snodprot1* homolog as a phytopathogenic factor.

Introduction

High throughput analyses are essential in functional genomics. The function of genes can be assessed through over-expression of target genes, misscheduled expression, and targeted disruption. Recently, gene silencing using the RNAi approach has been used extensively in mammalian systems. In contrast to the knockout deletion strategy, the RNAi knockdown strategy requires careful target site selection since this greatly influence efficacy. RNAi has been used in the rice blast fungus as a tool to knock down the expression of several target proteins (Kadotani et al., 2003; Nakayashiki et al., 2005). However, the level of effectiveness varied. Thus, targeted gene deletion strategy remains the most reliable tool for the functional analysis of candidate genes. However, for most eukaryotes, gene deletion is not as facile as in budding yeast. In filamentous fungi, the rate of gene deletion by homologous double recombination ranges from less than 1% to 20% of marker selected transformants, depending on the length of flanking sequences. To increase the rate of homologous recombination, a strategy termed TAG-KO using long flanking sequences has been used with some success in a large insert DNA libraries (Hamer et al., 2001). The generation of knockout constructs is also cumbersome. In budding yeast, short flanking sequences (~40bp) can be used for gene targeting due to the high rate of homologous recombination, thus facilitating the process of gene deletion construct generation. An adaptmer strategy which employs PCR-based cloning free deletion amplicon generation was shown to be effective for gene deletion in yeast (Reid et al., 2002). A strategy that uses a split marker gene was developed in yeast for efficient generation of target gene deletion strains (Fairhead et al., 1996). Since antibiotic resistance is obtained only through homologous recombination of overlapping selectable

marker gene *in vivo* in split-marker system, it should reduce the number of transformants that need to be screened to identify those that have undergone homologous recombination at the site of target gene. This strategy was tested in the rice blast fungus for gene deletion.

Snodprot1 is a member of a family of small secreted proteins identified in a number of fungal species (Skinner et al., 2001). The protein shows homology to a phytotoxic protein; its founding member is the biochemically characterized phytotoxic protein cerato-platanin (Pazzagli et al., 1999b), and contains the PFAM domain PF07249. Snodprot1 was first identified in *Phaeosphaeria (Septoria) nodorum* as a protein that is highly expressed during wheat infection (Hall et al., 1999). In the animal infecting fungus, *Aspergillus fumigatus*, the snodprot1 homolog is antigenic. Characterization of a snodprot1 homolog, CS-Ag, from the animal pathogen, *Coccidioides immitis*, showed that the protein is secreted and glycosylated (Pan and Cole, 1995). Interestingly, CS-Ag protein has been suggested to have serine proteinase activity. A homolog in *Neurospora crassa* was identified as a clock-controlled gene (Zhu et al., 2001). Similar light controlled expression was observed in *Leptosphaeria maculans*, the blackleg pathogen in *Brassica napus* (Wilson et al., 2002). Purified cerato-platanin protein from *Ceratocystis fimbriata* f. sp. *platani* caused autofluorescence and necrosis on tobacco leaves. Treatment of bacterially expressed and purified SP1 protein from *L. maculans* caused autofluorescence on *B. napus* leaves. Thus, members of this protein family appear to be phytotoxic in general. The spacing of Cys residues and similar motifs found in cerato-platanin and certo-ulmin led to the suggestion that the protein family may constitute a new family of hydrophobins. However, unlike typical hydrophobins, which have 8

cysteine residues that participate in intramolecular disulfide bond formation, the ceratoplatenin and snodprot1 protein family have 4 cysteine residues (Wosten, 2001). Although these proteins have been identified in a variety of ascomycetous fungi, functional analysis has been only recently investigated, and SP1 from *L. maculans* has a clearly identifiable signal peptide at the N-terminus, suggesting the protein is secreted (Wilson et al., 2002). Protein homologs were detected also in culture filtrate. Recent immunological evidence suggested that the protein was found on the cell wall, consistent with its possible role as a hydrophobin (Boddi et al., 2004).

In the work presented here, we knocked out the snodprot1 homolog in the rice blast fungus using the split marker strategy. Our results showed that split marker system is more efficient in generating deletion strains than conventional approaches. Analysis of two independent deletion strains revealed no apparent effect on vegetative growth and appressorium formation. However, the deletion mutants exhibited reduced pathogenicity on susceptible host plants. Epitope-tagged protein expression suggested no post-translational modification, and the protein was mainly detected in the culture filtrate.

Materials and methods

Culture, gene identification, and sequence analysis

Magnaporthe grisea 70-15 strain was used for this study. Cultures were grown on Oatmeal agar (OMA) for routine propagation under constant light. Liquid cultures were grown in synthetic complete medium (CM; 6 g yeast extracts, 6 g casein, 10 g sucrose, and 1 ml of *A. nidulans* trace element solution per liter). For expression analysis, culture

was grown in synthetic minimal medium (MM; Sucrose 10 g, *A. nidulans* trace element solution 1 ml, 20x nitrate salts 50 mL, thiamine 1mg, and biotin 5 µg per liter). The 20x nitrate salt solution was NaNO₃ 60 g, KCl 5.2 g, MgSO₄7H₂O 5.2 g, KH₂PO₄ 15.2 g in 500 ml. Cultures were grown with shaking in the dark. The snodprot1 homolog was identified from *M. grisea* proteins sequence (MG05344) using *N. crassa* snodprot1 homolog as a query. We named the rice blast homolog as MgSPH1 (*Magnaporthe grisea* Snodprot1 Homolog 1). BLAST analysis was done using MgSPH1 protein sequence to retrieve other fungal homologs (Altschul et al., 1997), and retrieved 13 protein sequences. A multiple sequence alignment was generated using ClustalX (Jeanmougin et al., 1998). The alignment was used for phylogenetic analysis using Neighbor joining tree algorithm in MEGA3 software (Kumar et al., 2004). The sequence logo was generated using Weblogo (Crooks et al., 2004). Sequence logo is a graphical representation of conserved amino acids or nucleotide using multiple sequence alignment. The hydrophobicity plot was generated using the EMBOSS package (Rice et al., 2000). The hydrophobin protein sequences and alignment were downloaded from the PFAM database (accession number: PF01185), and used for hydropathy plot analysis (Bateman et al., 2004). Hydrophobin data consisted of 30 protein sequences.

Gene deletion by split marker

The split-marker strategy is depicted in Figure 5.1. The deletion amplicon was generated as follows: Primer sequences were modified from the yeast protocol (Reid et al., 2002). The hygromycin cassette was amplified using primers AdHPHAscI (5'-CCGCTGCTAGGCGCGCCTAGCTGGAGCTAGTGGAGGTCA-3') and AdHPHFseI

(5'-GCAGGGATGGCCGGCCTTGGTCGGCATCTACTCTATTC-3') from pCB1004 (Carroll et al., 1994). Left flanking sequence was generated using SPL (5'-GCCATGGACAAGTTGGTCG-3') and SPLAd (5'-AGGCGCGCCTAGCTGCGGCA GAGGAGGCGATTTTGGC-3'). The right flanking sequence was amplified using SPR (5'-CTTACCCTCGGACCTTCGG-3') and SPRAd (5'-AGGCCGGCCATCCCTGCGGTGGTATTAGAGCGAGAGGC-3'). The two PCR fragments were about 500 bp. After amplification of each component, approximately 10 ng of left flank, right flank, and the hygromycin cassette was mixed in a 50 µl reaction with the SPL and SPR primers (0.5µM final concentration for each). The PCR conditions for adaptamer fusion PCR was 94 °C for 1 min followed by 30 cycles of 94 °C for 30 sec, 55 °C for 30 sec, and 68 °C for 3 min with a final extension at 68 °C for 5 min. To amplify the split marker amplicons, SLP::ΔC-HYG was amplified using HY (5'-GGATGCCTCCGCTCGAAGTA-3') and SPL primers, SPR::ΔN-HYG was generated using YG (5'-CGTTGCAAGACCTGCCT GAA-3') and SPR primers. Transformation was done using approximately 2 µg of PCR products without purification (Leung et al., 1990). Transformants were selected in regeneration medium with 400 µg/ml of hygromycin B. Transformants growing on the selection medium were transferred to fresh complete medium amended with 400 µg/ml of hygromycin B. Transformants were screened by PCR using the SPL and SPR primers. Gene knockouts were expected to produce a PCR product of approximately 2.4 kb compared to 1.2 kb for ectopic mutants. Integration of the hygromycin cassette was confirmed by PCR using hygromycin specific primers outside HY and YG primer binding sites described elsewhere (Xu and Hamer, 1996) .

Appressorium and pathogenicity assays

The appressorium assay was performed as described elsewhere (Choi and Dean, 1997; Mitchell and Dean, 1995). Pathogenicity of deletion strains was tested as follows. Two week old rice and two week old barley plant were spray inoculated with 2 ml of a conidial suspension adjusted to 1×10^5 /ml in 0.25% gelatin. Inoculated plants were incubated in a light regulated growth chamber with high humidity. Disease symptoms were rated one and two weeks after inoculation. The assay was repeated at least three times. Wound assays were performed using detached barley leaves. Barley leaves were wounded with sterile pipette tips, and 20 μ l of spore suspension was applied. Symptom development was monitored, and rated a week after inoculation.

Expression studies

RT-PCR analysis was performed on RNA isolated from following tissues. Conidial suspension was inoculated in liquid CM for 48 hrs. The fungal mycelium was collected by vacuum filtration, and was divided into liquid MM, MM-C, and MM-N, where carbon (sucrose) and nitrogen sources (NaNO_3) were omitted, respectively. The culture was grown for an additional 6 hrs, and mycelia were collected using vacuum filtration. RNA was prepared using TRIZOL reagent (Invitrogen). RNA samples were also prepared from cells grown on cellophane film for 8 hrs to induce appressoria formation. For the reverse transcription reaction, 5 μ g of total RNA was used in a 20 μ l reaction using Superscript III RT (Invitrogen). Two microliters of the RT mix was used for quantitative RT-PCR in a 50 μ l volume using primers snodprot1F (AAAAAAGCAGGCTTCACCATGCAGTTCTCCAACATCC) and snodprot1R (GTA

CAAGAAAGCTGGGTCCAGGCCGCAGGCGTTGAG). PCR condition was 94 °C for 1 min for initial denaturation followed by 26 cycles of 94 °C for 30 sec, 53 °C for 30 sec, and 72 °C for 1 min. And, final extension was at 72 °C for 3 min. PCR products were analyzed on a regular agarose gel.

MgSPH1 over-expression strain was a kind gift from Daniel Ebbole at Texas A&M University. The strain harbors the RGSHHHHHH epitope tag at the C-terminus of MgSPH1. Gene expression was under the control of RP27 promoter (Bourett et al., 2002). The strain was grown in CM liquid culture for 3 days, and the mycelial mass was removed by vacuum filtration. Proteins were purified from the culture filtrate using Ni-NTA beads according to the manufacturer's instructions (Qiagen). Purified protein was resolved on a 4-20% SDS-PAGE gel, and blotted to a PVDF membrane. Western analysis was performed using an anti-RGSHis antibody (Qiagen) and anti-mouse antibody alkaline phosphatase conjugate (Sigma). Signal was detected either using chromogenic BCIP/NBT solution (Sigma) or chemiluminescence (Roche).

Results

Sequence characterization of MgSPH1

A snodprot1 homolog MgSPH1 in the rice blast fungus was identified as MG05344.4 in the *M. grisea* genome database by BLAST using *N. crassa* snodprot1 homolog (GenBank accession number: CAC28585) as a query (<http://www.broad.mit.edu/annotation/fungi/magnaporthe/>) with the GenBank accession number XP_359969.1. BLASTP was run using the MgSPH1 predicted protein sequence.

BLAST analysis retrieved homologous sequences in a variety of fungi with the two closest hits being a probable snodprot1 homolog precursor in *Neurospora crassa* (CAC28585) with e-value of $2e-52$ and hypothetical protein FG11205.1 in *Gibberella zeae* (EAA75415) with e-value of $2e-47$. Other matches were to proteins from animal pathogens such as the allergen Asp13 from *Aspergillus fumigatus* Af293 (EAL93557), CS-antigen from *Coccidioides posadasii* (AAN73410). Also, cerato-platanin, a phytotoxic protein in *Ceratocystis fimbriata* f. sp. *platani* (CAC84090) showed a significant homology ($2e-17$). Despite being linked to phytotoxicity, snodprot1 homologs are present in saprophytes such as *N. crassa*. The homolog in *N. crassa* was identified as a clock-controlled gene (Zhu et al., 2001). A similar light-regulated expression pattern was also noted for the snodprot1 homolog in *Leptosphaeria maculans* (Wilson et al., 2002), suggesting a general transcription pattern/profile in filamentous fungi. As shown in Figure 5.2A, sequence alignment of MgSPH1 homologs suggested that four Cys residues were conserved, and might be important for proper confirmation through the formation of intramolecular disulfide bonds. Also, the phylogenetic analysis showed that homologs in *N. crassa* and *Fusarium* spp. were mostly related to MgSPH1 (Figure 5.2B). Pazzagli et al noted that cerato-platanin and cerato-ulmin may have a conserved signature of Cys-Ser-Asn and Cys-Cys-Asn, respectively (Pazzagli et al., 1999a). Our analysis indicated that Cys-Ser-Asp was prevalent over Cys-Ser-Asn in the protein sequences examined (Figure 5.2A). Thus, the biological importance of the signature remains to be determined. Hydropathy plots of fungal hydrophobins and MgSPH1 homologs from sequence alignment using pepwindowall software in EMBOSS

package revealed that overall hydrophobicity was higher in the hydrophobin protein family than for MgSPH1 homologs (Figure 5.3).

Split-marker was efficient in *in vivo* gene deletion

In many filamentous fungal systems, the generation of knockout strains is a major bottle neck for analyzing the function of a gene under investigation. Gene deletion by homologous recombination within the genome has been the traditional strategy. In this study, we developed a split marker system to replace the *snodprot1* homolog in the rice blast fungus genome with a hygromycin expression cassette, and compared the efficiency of the method with a classical strategy using the intact hygromycin cassette.

Results are summarized in Table 5.1. First of all, the number of transformants arising from split marker was smaller compared to the classical method. In a typical experiment, we were able to obtain 50~100 hygromycin resistant transformants from the transformation with intact hygromycin expression construct. In contrast, we observed 5-20 hygromycin resistant colonies from split marker transformation, suggesting homologous recombination between the overlapping marker sequences and flanking sequences occurred less frequently. Two batches of transformations were done for split marker transformation. From 25 transformants obtained by split marker, we identified 2 transformants that had undergone gene replacement of MgSPH1. In contrast, no deletion strains were obtained from analysis of >120 transformants obtained using the classical strategy. The deletion strains were named as Δ SPH1-1 and Δ SPH1-2. An ectopic strain was selected as a control, and was designated as SPect. As shown in Figure 5.4A, PCR analysis of wild type and hygromycin-resistant strain indicated that MgSPH1 was

replaced by hygromycin cassette in Δ SPH1-1 and Δ SPH1-2, and SPect had ectopic integration of hygromycin cassette. As expected, two bands were observed in SP3. Integration of hygromycin cassette was confirmed by hygromycin primers as described in materials and methods (Figure 5.4B). These strains along with the wild type recipient strain were selected for further analysis. Our data strongly suggests that the split marker strategy is a useful tool for homologous recombination mediated gene deletion in the rice blast fungus.

MgSPH1 is dispensable for appressorium formation

Growth characteristics of deletion strains indicated that one of the strains, Δ SPH1-1, showed a slightly reduced growth and sporulation on OMA while Δ SPH1-2 was indistinguishable from wild type and ectopic integration strain, SPect. Even though the exact role of the MgSPH1 homologs in many fungi is unknown, it has been suggested that the protein might be a hydrophobin. Functional analysis of MPG1, a hydrophobin in *M. grisea* supports a role in surface sensing. Deletion of the hydrophobin gene MPH1 in *M. grisea* manifests pleiotropic effects, including decreased conidiation and *in planta* growth (Kim et al., 2005). Deletion of MgSPH1 did not affect the ability to form appressoria on hydrophobic surface of GelBond. Also, the mutants formed appressoria indistinguishable from wild type in response to inducers such as cAMP, 1,16-hexdecane diol, and IBMX (data not shown). Examination of inoculated barley leaves revealed the presence of normal looking appressoria (data not shown). Thus, MgSPH1 is dispensable for appressorium formation under the conditions we investigated.

MgSPH1 is required for pathogenicity

Work with *snodprot1* homologs in other fungi suggested that this protein may play a role in disease. To investigate the role of MgSPH1 in pathogenicity, we spray inoculated young rice and barley plants. Evaluation of plants one week after the inoculation revealed that deletion mutants Δ SPH1-1 and Δ SPH1-2 were greatly reduced in their ability to cause disease (Figure 5.5A). The ectopic mutant SPect and wild type strains showed a similar degree of disease development in both rice and barley. Both strains showed extensive and rapid development of disease symptoms. Both Δ SPH1-1 and Δ SPH1-2 mutants produced considerably fewer and smaller lesions compared to the wild type and SPect ectopic strain. To further investigate if the reduced pathogenicity was caused by a defect in *in planta* growth, we performed wounded leaf assays (Figure 5.5B). Barley leaves were wounded with sterile pipette tips, and inoculated with 20 μ l of spore suspension (10^5 conidia/ml). Again, the ectopic strain SPect exhibited a similar degree of symptom development as the wild type. However, the deletion mutants Δ SPH1-1 and Δ SPH1-2 showed little development of disease, suggesting that MgSPH1 is required for invasive *in planta* growth.

MgSPH1 recombinant protein does not appear to be phytotoxic

To further investigate the role of the protein in host-fungal pathogen interaction, we purified RGSHHHHHHH epitope tagged MgSPH1 under the control of constitutive promoter in the rice blast fungus grown in CM. The protein was mainly detected in the culture filtrate. SDS-PAGE and Western analysis are shown in Figure 5.6. The MW of epitope-tagged MgSPH1p is 12~13 KDa, very similar to the predicted MW of mature

MgSPH1, suggesting that the protein was not glycosylated. This contrasts with the homolog in *C. immitis*, which was shown to be glycosylated. We were not able to detect signal following digestion of the cell wall with β -glucanase, suggesting that MgSPH1 is not associated with cell wall under the conditions we used.

For phytotoxicity assays, approximately 400 ng of purified and de-salted protein was injected into wound sites on detached barley leaves. As a control, wild type culture filtrate was processed the same way as the filtrate of MgSPH1 over-expression strain. Also, 400 ng of BSA and water were added as controls. In contrast to the phytotoxicity exhibited by purified cerato-platanin from *C. fimbriata* f. sp. *platani* and SP1 from *L. maculans*, no apparent necrosis was visible on barley after 48 hrs inoculation. In this case, 30 ng of purified SP1 from *L. maculans* was sufficient to cause a plant response. However, SP1 from *S. nodorum* did not show any biological activity, including protease activity and phytotoxicity (Hall et al., 1999). Thus, the mechanism of action of MgSPH1 during the pathogenic interaction remains to be determined.

Expression of MgSPH1 is sensitive to nutrient starvation

We investigated the expression of MgSPH1 using RNA samples from different conditions. As shown in Figure 5.7, we observed increased transcription levels during nitrogen deprivation compared to growth in synthetic minimal medium and carbon deprivation. RT-PCR analysis also indicated that MgSPH1 was not up-regulated during appressorium formation under the conditions we investigated, consistent with its dispensable role of MgSPH1 during appressorium formation. Elevated expression levels of MgSPH1 during nitrogen-starvation and no changes in expression during appressorium

compared to mycelium was confirmed by whole genome microarray analysis (unpublished data). This expression pattern is consistent with its probable role in *in planta* growth during colonization.

Discussion

In this work, we characterized MgSPH1, a snodprot1 homolog in *M. grisea*. MgSPH1 shows homology to cerato-platanin, a phytotoxic protein found in *C. fimbriata* f. sp. *platani*. Further analysis revealed homologs in *N. crassa* as well as in other animal and plant pathogenic fungi. To date, the role of this protein remains elusive. However, all homologs harbor a signal peptide, suggesting an extracellular function. In some cases, secretion has been demonstrated. Amino acids alignment analysis revealed that Cys residues, which are likely to be involved in intramolecular disulfide bonds formation, appear to be highly conserved, suggesting the formation of disulfide bond might be essential for function. For example, MPG1, a hydrophobin gene in the rice blast fungus, has eight cysteine residues, all of which participate in intramolecular disulfide bonds. The substitution of Cys->Ala in MPG1 resulted in a defect in secretion and cell wall localization in *M. grisea* (Kershaw et al., 2005). Pazzagli et al. suggested that Cys-Ser-Asn in cerato-platanin and Cys-Cys-Asn in cerato-ulmin might constitute a hydrophobin signature (Boddi et al., 2004; Pazzagli et al., 1999b). Sequence alignment of fungal snodprot1 homologs and the inspection of the sequence logo indicated that Asn was predominant over Asp. Currently, it is not known well whether the signature sequence is important for function. Interestingly, unlike fungal hydrophobins such as MPG1, which are highly divergent, BLAST analysis of the snodprot1 homologs retrieved thirteen

putative homologs, suggesting the sequence divergence is more restricted compared to authentic hydrophobins. The stringent sequence conservation suggests that the role of snodprot1 homologs might be more specific than hydrophobins.

Functional analysis of MgSPH1 from *M. grisea* contrasts with observation made for SP1, the snodprot1 homolog in *L. maculans*. Deletion of SP1 in *L. maculans* had no effect on pathogenicity. Also, localization studies suggested that the homologous protein in other fungi localizes to the fungal cell wall. However, in our work, MgSPH1 protein was primarily detected in the culture filtrate. A cell wall digest by β -glucanase did not reveal detectable amounts of MgSPH1 protein, suggesting MgSPH1p is secreted into medium and does not associate to cell wall. Hydrophobin assembly requires a hydrophobic-hydrophilic interface or water-air interface (Wessels, 1994). McCabe et al. suggested that Cryparin, a fungal hydrophobin from *Cryphonectria parasitica*, is secreted into medium, binds to, and remains on the cell wall in submerged culture (McCabe and Van Alfen, 1999). There is no report on self-assembly of snodprot1 homologs as reported for other hydrophobins. MHP1, a class II hydrophobin in *M. grisea*, was isolated from cDNA library prepared from infected rice (Kim et al., 2005). Deletion of MHP1 resulted in impaired *in planta* growth. Also, studies of SP1 from *S. nodorum* indicated that mRNA expression and protein increased during tissue colonization. However, these results may simply represent increase in fungal biomass rather than elevated expression. Many pathogenicity determinants are known to be nutrient responsive (Lau and Hamer, 1996; Van den Ackerveken et al., 1994). Expression of MgSPH1 showed that its expression is also nutrient-sensitive and induced upon nitrogen starvation. Nitrogen starvation conditions have been proposed to mimic the *in planta* environment. Thus, it appears that

MgSPH1 might be required for efficient growth *in planta*. However, no apparent enzymatic activity or biological function of the *snodprot1* homologs has been identified with the exception of Cs-Ag in *C. immitis*.

In this work, we developed and demonstrated the utility of the split-marker knockout strategy for functional analysis of gene in *M. grisea*. Using a relatively short stretch of flanking sequences, it was possible to recover homologous recombinant transformants. Other strategies such as positive-negative selection using a suicide gene have been devised to increase the rate of homologous recombination (Khang et al., 2005). Such a strategy offers an attractive advantage over traditional single dominant marker system. However, the method requires construction of transformation vector. In the method we developed for *M. grisea*, the transforming PCR product can be easily generated by two step PCR reaction either using ligation based fusion PCR or recombination fusion PCR strategy, obviating the need for the cloning step. Thus, this method offers considerable potential for high throughput manner for generating PCR products for use in transformation. We recognize that although the strategy yielded an obvious increase in deletion mutant generation, it is possible that some genes, potentially essential genes, might be recalcitrant to homologous recombination.

In conclusion, we applied the split-marker system for a gene replacement in the rice blast fungus. The split-marker system resulted in at least a 10-fold increase in rate of homologous recombination compared to the fusion PCR strategy using an intact hygromycin cassette. This suggests that split-marker can be generally applied to the rice blast fungus as a viable strategy for gene deletion. This strategy was used to characterize

MgSHP1, a *snodprot1* homolog. We demonstrated that MgSPH1 is an important pathogenicity gene in *M. grisea*, and is required for growth *in planta*.

References

1. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W., and Lipman,D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402.
2. Bateman,A., Coin,L., Durbin,R., Finn,R.D., Hollich,V., Griffiths-Jones,S., Khanna,A., Marshall,M., Moxon,S., Sonnhammer,E.L.L., Studholme,D.J., Yeats,C., and Eddy,S.R. (2004). The Pfam protein families database. *Nucl. Acids Res.* 32, D138-D141.
3. Boddi,S., Comparini,C., Calamassi,R., Pazzagli,L., Cappugi,G., and Scala,A. (2004). Cerato-platanin protein is located in the cell walls of ascospores, conidia and hyphae of *Ceratocystis fimbriata* f. sp. *platani*. *FEMS Microbiology Letters* 233, 341-346.
4. Bourett,T.M., Sweigard,J.A., Czymmek,K.J., Carroll,A., and Howard,R.J. (2002). Reef coral fluorescent proteins for visualizing fungal pathogens. *Fungal Genet. Biol* 37, 211-220.
5. Carroll,A.M., Sweigard,J.A., and Valent,B. (1994). Improved vectors for selecting resistance to hygromycin. *Fungal Genet Newslett* 41, 22.

6. Choi,W. and Dean,R.A. (1997). The adenylate cyclase gene MAC1 of *Magnaporthe grisea* controls appressorium formation and other aspects of growth and development. *Plant Cell* 9, 1973-1983.
7. Crooks,G.E., Hon,G., Chandonia,J.M., and Brenner,S.E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188-1190.
8. Fairhead,C., Llorente,B., Denis,F., Soler,M., and Dujon,B. (1996). New vectors for combinatorial deletions in yeast chromosomes and for gap-repair cloning using 'split-marker' recombination. *Yeast* 12, 1439-1457.
9. Hall,N., Keon,J., and Hargreaves,J. (1999). A homologue of a gene implicated in the virulence of human fungal diseases is present in a plant fungal pathogen and is expressed during infection. *Physiological and Molecular Plant Pathology* 55, 69-73.
10. Hamer,L., Adachi,K., Montenegro-Chamorro,M.V., Tanzer,M.M., Mahanty,S.K., Lo,C., Tarpey,R.W., Skalchunes,A.R., Heiniger,R.W., Frank,S.A., Darveaux,B.A., Lampe,D.J., Slater,T.M., Ramamurthy,L., DeZwaan,T.M., Nelson,G.H., Shuster,J.R., Woessner,J., and Hamer,J.E. (2001). Gene discovery and gene function assignment in filamentous fungi. *Proc. Natl Acad. Sci U. S. A* 98, 5110-5115.
11. Jeanmougin,F., Thompson,J.D., Gouy,M., Higgins,D.G., and Gibson,T.J. (1998). Multiple sequence alignment with Clustal X. *Trends Biochem Sci* 23, 403-405.

12. Kadotani,N., Nakayashiki,H., Tosa,Y., and Mayama,S. (2003). RNA silencing in the phytopathogenic fungus *Magnaporthe oryzae*. *Mol Plant Microbe Interact.* *16*, 769-776.
13. Kershaw,M.J., Thornton,C.R., Wakley,G.E., and Talbot,N.J. (2005). Four conserved intramolecular disulphide linkages are required for secretion and cell wall localization of a hydrophobin during fungal morphogenesis. *Mol Microbiol* *56*, 117-125.
14. Khang,C.H., Park,S.Y., Lee,Y.H., and Kang,S. (2005). A dual selection based, targeted gene replacement tool for *Magnaporthe grisea* and *Fusarium oxysporum*. *Fungal Genet. Biol* *42*, 483-492.
15. Kim,S., Ahn,I.P., Rho,H.S., and Lee,Y.H. (2005). MHP1, a *Magnaporthe grisea* hydrophobin gene, is required for fungal development and plant colonization. *Mol Microbiol* *57*, 1224-1237.
16. Kumar,S., Tamura,K., and Nei,M. (2004). MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief. Bioinform.* *5*, 150-163.
17. Lau,G. and Hamer,J.E. (1996). Regulatory genes controlling MPG1 expression and pathogenicity in the rice blast fungus *Magnaporthe grisea*. *Plant Cell* *8*, 771-781.

18. Leung,H., Lehtinen,U., Karjalainen,R., Skinner,D., Tooley,P., Leong,S., and Ellingboe,A. (1990). Transformation of the rice blast fungus *Magnaporthe grisea* to hygromycin B resistance. *Curr. Genet.* 17, 409-411.
19. McCabe,P.M. and Van Alfen,N.K. (1999). Secretion of cryparin, a fungal hydrophobin. *Appl. Environ. Microbiol* 65, 5431-5435.
20. Mitchell,T.K. and Dean,R.A. (1995). The cAMP-dependent protein kinase catalytic subunit is required for appressorium formation and pathogenesis by the rice blast pathogen *Magnaporthe grisea*. *Plant Cell* 7, 1869-1878.
21. Nakayashiki,H., Hanada,S., Nguyen,B.Q., Kadotani,N., Tosa,Y., and Mayama,S. (2005). RNA silencing as a tool for exploring gene function in ascomycete fungi. *Fungal Genet Biol* 42, 275-283.
22. Pan,S. and Cole,G.T. (1995). Molecular and biochemical characterization of a *Coccidioides immitis*- specific antigen. *Infect. Immun.* 63, 3994-4002.
23. Pazzagli,L., Cappugi,G., Manao,G., Camici,G., Santini,A., and Scala,A. (1999a). Purification, characterization, and amino acid sequence of cerato- platanin, a new phytotoxic protein from *Ceratocystis fimbriata* f. sp. *platani*. *J. Biol. Chem.* 274, 24959-24964.

24. Pazzagli,L., Cappugi,G., Manao,G., Camici,G., Santini,A., and Scala,A. (1999b). Purification, characterization, and amino acid sequence of cerato-platanin, a new phytotoxic protein from *Ceratocystis fimbriata* f. sp. *platani*. *J. Biol. Chem.* *274*, 24959-24964.
25. Reid,R.J., Lisby,M., and Rothstein,R. (2002). Cloning-free genome alterations in *Saccharomyces cerevisiae* using adaptamer-mediated PCR. *Methods Enzymol.* *350*, 258-277.
26. Rice,P., Longden,I., and Bleasby,A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* *16*, 276-277.
27. Skinner,W., Keon,J., and Hargreaves,J. (2001). Gene information for fungal plant pathogens from expressed sequences. *Current Opinion in Microbiology* *4*, 381-386.
28. Van den Ackerveken,G.F., Dunn,R.M., Cozijnsen,A.J., Vossen,J.P., Van den Broek,H.W., and de Wit,P.J. (1994). Nitrogen limitation induces expression of the avirulence gene *avr9* in the tomato pathogen *Cladosporium fulvum*. *Mol Gen Genet* *243*, 277-285.
29. Wessels,J.G.H. (1994). Developmental regulation of fungal cell wall formation. *Annu. Rev Phytopathol.* *32*, 413-437.

30. Wilson,L.M., Idnurm,A., and Howlett,B.J. (2002). Characterization of a gene (sp1) encoding a secreted protein from *Leptosphaeria maculans*, the blackleg pathogen of *Brassica napus*. *Molecular Plant Pathology* 3, 487-493.
31. Wosten,H.A. (2001). Hydrophobins: multipurpose proteins. *Annu. Rev Microbiol* 55, 625-646.
32. Xu,J.R. and Hamer,J.E. (1996). MAP kinase and cAMP signalling regulate infection structure formation and pathogenic growth in the rice blast fungus *Magnaporthe grisea*. *Genes Dev.* 10, 2696-2706.
33. Zhu,H., Nowrousian,M., Kupfer,D., Colot,H.V., Berrocal-Tito,G., Lai,H., Bell-Pedersen,D., Roe,B.A., Loros,J.J., and Dunlap,J.C. (2001). Analysis of expressed sequence tags from two starvation, time-of-day-specific libraries of *Neurospora crassa* reveals novel clock-controlled genes. *Genetics* 157, 1057-1065.

Table 5.1 Screening of transformants for homologous recombination

	split marker		Whole construct
	#1 ^a	#2	
# transformants screened ^b	20	5	>120
# deletion generated ^c	1	1	0
% deletion strains	5%	20%	0

^a Two transformation experiments were conducted using split-marker (#1 and #2), and one experiment using intact hygromycin expression cassette (whole construct).

^b Indicates number of transformants screened by PCR.

^c Indicates number of transformants in which MgSPH1 was deleted.

Figure legends

Figure 5.1 The strategy for the deletion of MgSPH1 by split-marker gene replacement system.

Genomic regions of left flanking (LF) and right flanking (RF) sequences of MgSPH1 were PCR amplified. The chimeric PCR products corresponding to LF:: Δ C-hyg and Δ N-hyg::RF were amplified using SPL and HY primer pair and SPR and YG primer pair, respectively. The resulting PCR products were mixed together for transformation into *M. grisea*.

Figure 5.2 Sequence and phylogenetic analyses of MgSPH1 homologs.

Sequence alignment was performed using Clustalx software from which a sequence logo was generated (panel (A)) using sequences listed in panel (B). The alignment of the sequences suggests that 4 cysteine residues (marked by asterisks) are conserved and may participate in disulfide bond formation. Also, the sequence Cys-Ser-(Asn/Asp) is highlighted by dashed line on top. A phylogenetic tree was generated using the Neighbor-joining algorithm (panel (B)). Bootstrap values are indicated at each node.

Figure 5.3 Kyte-Doolittle hydropathy plots of hydrophobins and MgSPH1 homologs.

Thirty hydrophobin protein sequences and alignment downloaded from PFAM database (PF01185) were used for the generation of hydropathy profile (panel A). Thirteen snodprot1 homologs (shown in Figure 2) identified by BLAST search were used to create hydropathy profile (panel B). Arrows indicate positions of conserved cysteine residues. The comparison of the profile suggests that MgSPH1 homologs exhibit less hydrophobicity compared to hydrophobin proteins.

Figure 5.4 Deletion of MgSPH1 was confirmed by PCR.

The deletion of MgSPH1 was confirmed by PCR analysis. (A) For identification of deletion mutants, SPL and SPR primers were used to screen transformants. The PCR product (2.2 kb) in which MgSPH1 ORF was replaced by hygromycin-resistance cassette indicates legitimate transformants while WT and ectopic transformants had a 1.1 kb PCR product. The PCR analysis using primers immediately flanking the region showed two bands in SPect strain, suggesting recombination of hygromycin cassette and loss of SPL and SPR primer binding sites during ectopic integration (data not shown). M: marker (λ HindIII), WT: wild type 70-15, Δ SPH1-1 and Δ SPH1-2: MgSP1 deletion strains, SPect: ectopic strain. (B) PCR analysis to confirm the integration of hygromycin marker in the transformants.

Figure 5.5 MgSPH1 is required for fungal pathogenicity.

(A) Spore suspension of wild type, two deletion strains (Δ SPH1-1 and Δ SPH1-2), and ectopic strain SPect was sprayed on to two weeks old susceptible barley. Progression of disease was monitored two weeks after inoculation. Wild type and ectopic strains caused rapidly developing symptoms. However, plants inoculated with deletion strains showed limited lesion development. (B) Detached barley leaves were drop inoculated with spore suspension with or without wounding with sterile tips. Wounding treatment is indicated by (w) following strain identifier. In contrast to the clearly identifiable lesion produced by wild type and SPect, Δ SPH1-1 and Δ SPH1-2 exhibited limited lesion development.

Figure 5.6 Purification of recombinant MgSPH1 protein.

Recombinant MgSPH1p with RGSHHHHHH tag was purified using Ni-NTA agarose beads from culture filtrate. Purified recombinant MgSPH1 protein was resolved on 4~20% SDS-PAGE (panel A). Apparent molecular weight was approximately 12~13 kDa, similar to the expected molecular weight 12.3 kDa. Panel (B) Western blot analysis of MgSPH1 performed on a SDS-PAGE gel. Samples are loaded as follows: Lane 1: purified MgSPH1p, lane 2: culture filtrate from MgSPH1 over-expression strain, lane 3: wild type culture filtrate, lane 4: cell wall digest, lane 5: mock (MgSPH1 overexpression strain incubated in 1 M sorbitol without addition of β -glucanase). Lower section of panel B shows Western blot probed with anti-RGSHis antibody. Western blot suggests that MgSPH1 is not associated with the cell wall.

Figure 5.7 Expression of MgSPH1 is nutrient-sensitive.

RT-PCR analysis was performed using RNA samples from different conditions. RNA samples from cultures grown in MM, MM-C, MM-N, and appressorial sample were subjected to RT-PCR. RT-PCR analysis showed that MgSPH1 was slightly up-regulated upon nitrogen starvation compared to synthetic minimal medium, carbon starvation and appressorial stage. MW: Molecular weight marker, M: synthetic minimal medium, -C: synthetic minimal medium without sucrose, -N: synthetic minimal medium without NaNO_3 as a nitrogen source, A: appressorium.

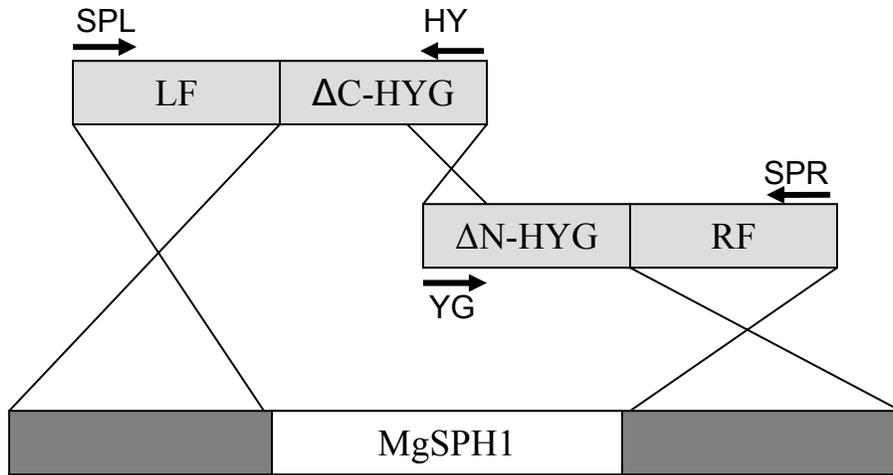


Figure 5.1 The strategy for the deletion of MgSPH1 by split-marker gene replacement system

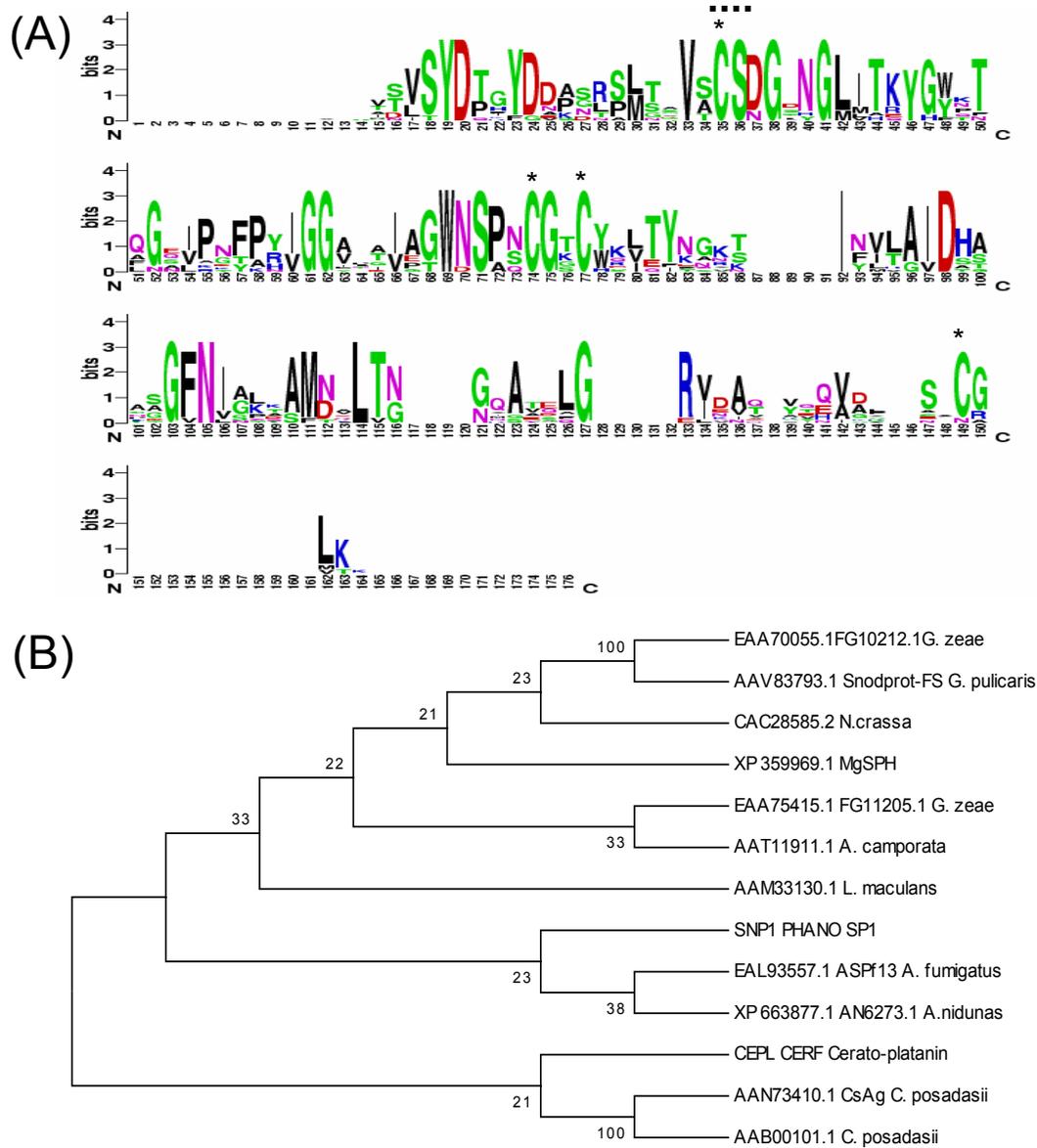
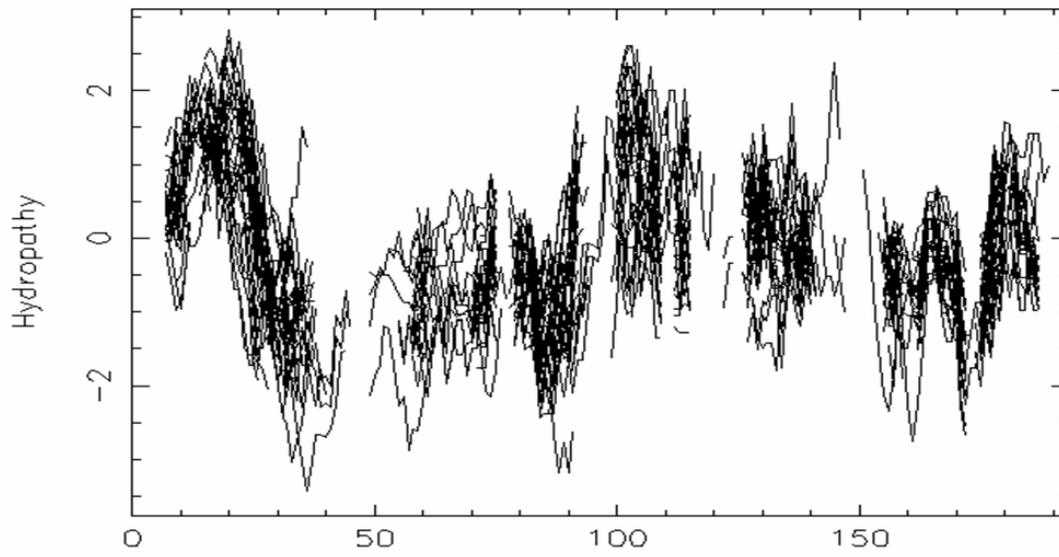


Figure 5.2 Sequence and phylogenetic analyses of MgSPH1 homologs

(A)



(B)

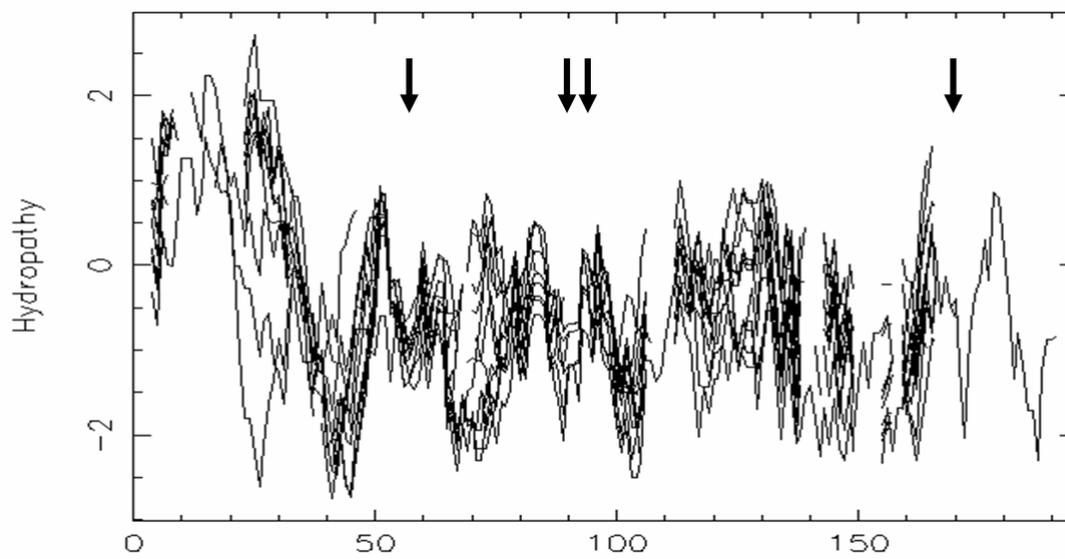


Figure 5.3 Kyte-Doolittle hydropathy plots of hydrophobins and MgSPH1 homologs

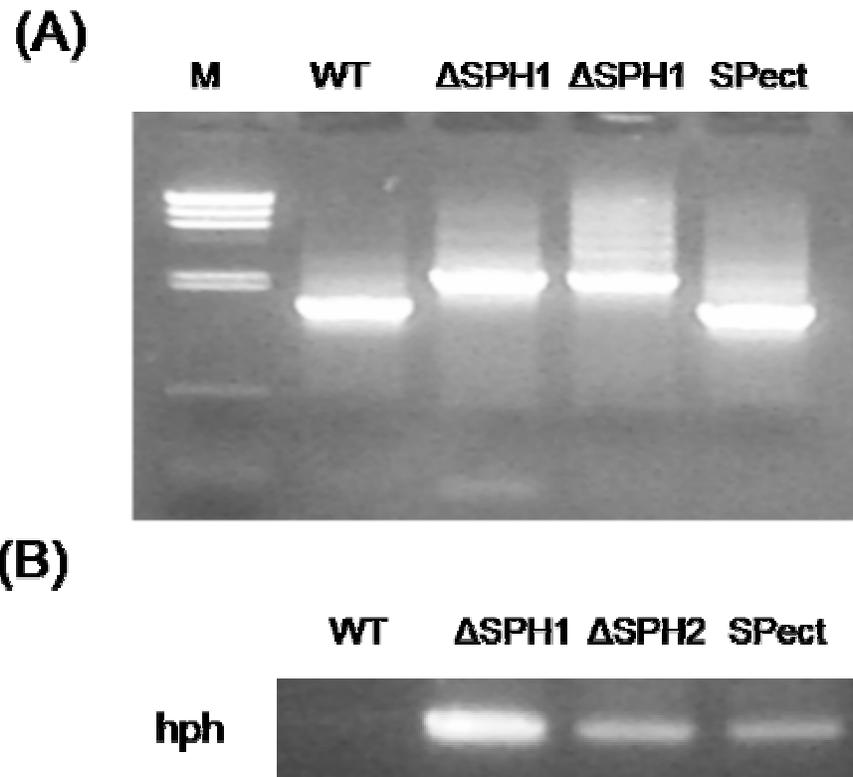
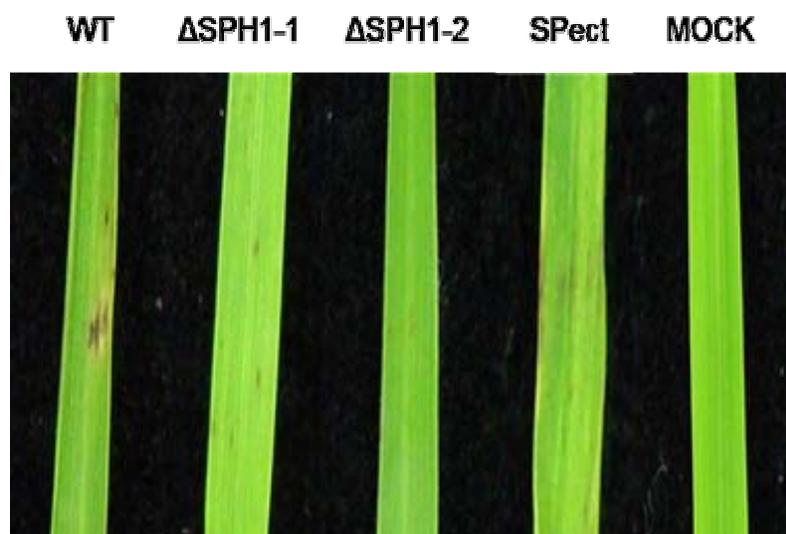


Figure 5.4 Deletion of MgSPH1 was confirmed by PCR

(A)



(B)

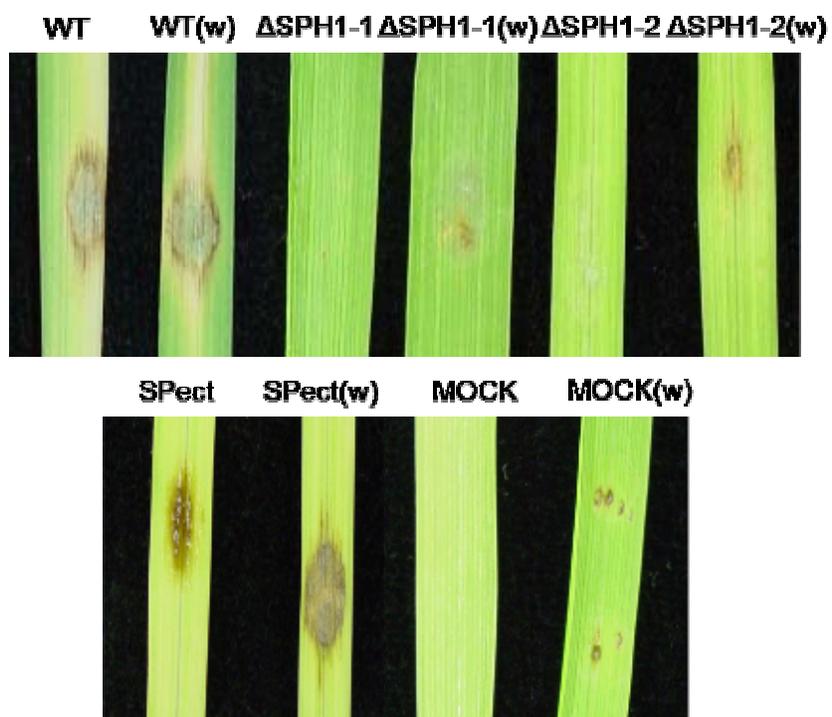
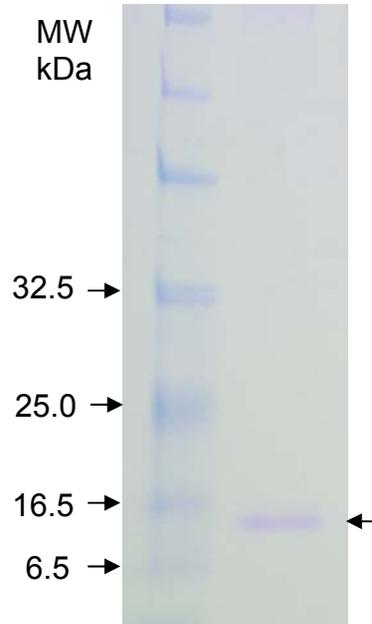


Figure 5.5 MgSPH1 is required for fungal pathogenicity

(A)



(B)

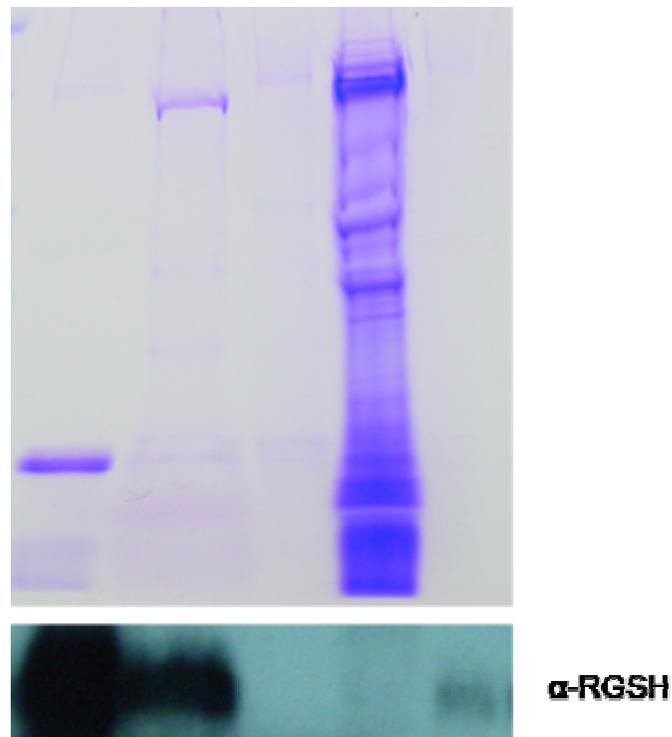


Figure 5.6 Purification of recombinant MgSPH1 protein

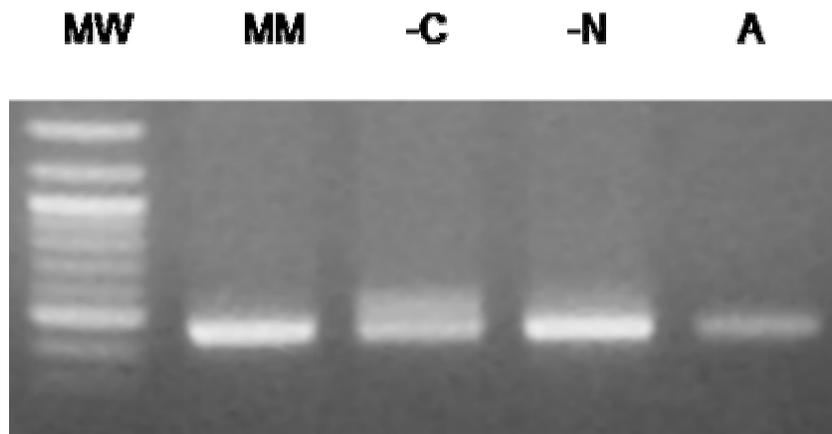


Figure 5.7 Expression of MgSPH1 is nutrient-sensitive

General Conclusion

In this work, we explored the functional genomics of the rice blast fungus, *M. grisea*. Rice blast disease causes enormous annual losses in rice production world-wide (Zeigler et al., 1994). Much research on *M. grisea* has focused on identifying genetic determinants of infection structure formation, revealing important signal transduction pathways (Dean, 1997; Xu, 2000). However, relatively little research has been performed regarding the post-penetration process. Moreover, recent work has indicated that the formation of appressorium may not be the prerequisite for host colonization of roots, which operates in gene-for-gene relationship (Sesma and Osbourn, 2004). This indicates that there may be a general mechanism of pathogenesis by plant pathogens whether they attack above or below ground, and highlights the importance of post-penetration studies. Work presented here in previous chapters indicates that post-penetration stages may represent important targets to thwart rice blast disease. A 14-3-3 homolog in *M. grisea*, MgFTT1, was characterized. Deletion of the MgFTT1 resulted in a greatly reduced pathogenicity, which was ascribed to impaired growth *in planta*. 14-3-3 proteins participate in a numerous signal transduction pathways where they bind phosphoproteins. Protein kinases which operate in signal transduction pathways are modulated by phosphorylation. Large scale proteomic and genomic studies have indicated that 14-3-3 binding partners function in diverse functions. Thus, proteomic approaches such as co-immunoprecipitation and yeast two hybrid analysis might lead to better understanding of *in vivo* role of 14-3-3. Also, elucidation of signal pathways involving MgFTT1 might unveil key components of post-penetration growth in *M. grisea*. In addition, relatively

unexplored is the crosstalk between signal pathways in the rice blast fungus. Biochemical studies are needed to provide a bridge to proteomic research. Proteomic studies will provide deeper insight into the pathogenic process since function of genes, in most part, is executed at the protein level. Thus, protein-protein interaction studies and identification of 14-3-3 binding partners will likely reveal a detailed picture of biochemical function exerted by MgFTT1 in pathogenesis. Deletion of a snodprot1 homolog, MgSPH1, resulted in reduced pathogenicity. Other studies conducted on the homologs of snodprot1 gene suggested that the transcription of snodprot1 homologs is under circadian control, hinting that snodprot1 homologs might play a developmental role. Its expression pattern suggested that transcript level is influenced by nutrient availability and growth stages. Currently, the exact physiological role of snodprot1 homologs is undetermined. The data presented in this work suggested that the gene might be required for efficient growth *in planta*.

Large scale EST sequencing has revealed a plethora of genetic information and dynamics of transcriptome in response to growth conditions in the rice blast fungus (Jantasuriyarat et al., 2005; Kim et al., 2001). Relatively unexplored is the post-penetration transcriptome. This might reflect difficulties in obtaining pure fungal biomass separated from plant tissues or detecting pathogen gene expression *in planta*. A recent study in the flax rust fungus, *Melampsora lini*, highlights the importance of post-penetration transcriptome studies, and suggested a way to obtain fungal biomass preparation during *in planta* growth (Catanzariti et al., 2006). Cytological studies have suggested that *in planta* growth may be different from *in vitro* growth. Thus, future transcriptome analysis of the post-penetration stages may be enlightening. The cDNA

library generation method presented in this work may be utilized for the generation of cDNA libraries targeted for post-penetration growth. Also, genome-wide microarray analysis of post-penetration transcriptome dynamics might reveal key physiological players in pathogenesis. However, this presents some technical issues to be resolved. First, microarray analysis of host-pathogen interaction requires a reference sample that can be compared with mixed biomass of pathogen and host tissues. As of now, this appears to be the main issue to be resolved. Also, limiting biomass of pathogen at the early stage of infection poses another problem, which directly affects the quality of microarray data. In this regard, the combination of laser capture microdissection (LCM) and gene expression profiling is attractive to address this issue, and may provide a more targeted approach (Rubin, 2001). Since only selected cells are culled from tissues under direct observation, LCM may provide a more precisely defined sample of tissues. Second, efficient detection of pathogen gene expression *in planta* might provide a better understanding of transcriptome dynamics through genomic tools such as serial analysis of gene expression (SAGE) (Velculescu et al., 1995; Irie et al., 2003), massively parallel signature sequencing (MPSS) (Brenner et al., 2000) or cDNA library generation and EST sequencing if the fungal tissue can be enriched as illustrated by the work in the *M. lini* and *Uromyces fabae* (Catanzariti et al., 2006; Hahn and Mendgen, 1997). Third, devising an *in vitro* system that mimics host tissues may provide an efficient way of obtaining appropriate tissue. Currently, efficient formation of infection structures by *M. grisea* occurs *in vitro* on a hydrophobic surface and by the application of chemical inducers (Lee and Dean, 1994; Lee and Dean, 1993). Finally, development of high throughput functional assay platforms is imperative. The genome-wide level functional genomic

study in budding yeast employed small sequence tags for identification of each deletion strains (Shoemaker et al., 1996). This provided a great advantage for *en mass* functional studies of deletion strains for fitness assessment in certain environments and sensitivity to toxic compounds, the latter might lead to the identification of novel target of antimycotic compounds using mutant pools. Also, efficient and parallel generation of expression constructs may be facilitated by the utilization of recombinational cloning in a variety of formats such as GATEWAY, cre-lox, and yeast gap-repair cloning systems, collectively termed as recombineering (Copeland et al., 2001).

In conclusion, with a wealth of genomic information and resources, *M. grisea* has emerged as a premier model organism to study fungal pathogen-host interaction. In contrast to the extensive research performed on the infection structure formation, relatively little attention has been paid to the interface of the interaction and the post-penetration growth. This apparent lack of attention may be due to technical experimental challenges such as difficulty of obtaining materials and assessment of expression. An integrated genomic approach will lead to a better understanding of the mechanisms of pathogenesis in the rice blast fungus.

References

1. Brenner,S., Johnson,M., Bridgham,J., Golda,G., Lloyd,D.H., Johnson,D., Luo,S., McCurdy,S., Foy,M., Ewan,M., Roth,R., George,D., Eletr,S., Albrecht,G., Vermaas,E., Williams,S.R., Moon,K., Burcham,T., Pallas,M., DuBridge,R.B., Kirchner,J., Fearon,K., Mao,J., and Corcoran,K. (2000). Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat Biotechnol.* *18*, 630-634.
2. Catanzariti,A.M., Dodds,P.N., Lawrence,G.J., Ayliffe,M.A., and Ellis,J.G. (2006). Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* *18*, 243-256.
3. Copeland,N.G., Jenkins,N.A., and Court DL (2001). Recombineering: a powerful new tool for mouse functional genomics. *Nat Rev Genet* *2*, 769-779.
4. Dean,R.A. (1997). Signal pathways and appressorium morphogenesis. *Annu. Rev. Phytopathol.* *35*, 211-234.
5. Hahn,M. and Mendgen,K. (1997). Characterization of in planta-induced rust genes isolated from a haustorium-specific cDNA library. *Mol Plant Microbe Interact.* *10*, 427-437.

6. Irie,T., Matsumura,H., Terauchi,R., and Saitoh,H. (2003). Serial Analysis of Gene Expression (SAGE) of *Magnaporthe grisea*: genes involved in appressorium formation. *Mol Genet Genomics* 270, 181-189.
7. Jantasuriyarat,C., Gowda,M., Haller,K., Hatfield,J., Lu,G., Stahlberg,E., Zhou,B., Li,H., Kim,H., Yu,Y., Dean,R.A., Wing,R.A., Soderlund,C., and Wang,G.L. (2005). Large-scale identification of expressed sequence tags involved in rice and rice blast fungus interaction. *Plant Physiol* 138, 105-115.
8. Kim,S., Ahn,I.P., and Lee,Y.H. (2001). Analysis of genes expressed during rice-*Magnaporthe grisea* interactions. *Mol Plant Microbe Interact.* 14, 1340-1346.
9. Lee,Y.H. and Dean,R.A. (1993). cAMP regulates infection structure formation in the plant pathogenic fungus *Magnaporthe grisea*. *Plant Cell* 5, 693-700.
10. Lee,Y.H. and Dean,R.A. (1994). Hydrophobicity of contact surface induces appressorium formation in *Magnaporthe grisea*. *FEMS Microbiol Lett.* 115, 71-75.
11. Rubin,M.A. (2001). Use of laser capture microdissection, cDNA microarrays, and tissue microarrays in advancing our understanding of prostate cancer. *J Pathol* 195, 80-86.
12. Sesma,A. and Osbourn,A.E. (2004). The rice blast pathogen undergoes developmental processes typical of root-infecting fungi. *Nature* 431, 582-586.

13. Shoemaker,D.D., Lashkari,D.A., Morris,D., Mittmann,M., and Davis,R.W. (1996). Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nat Genet* 14, 450-456.
14. Velculescu,V.E., Zhang,L., Vogelstein,B., and Kinzler,K.W. (1995). Serial analysis of gene expression. *Science* 270, 484-487.
15. Xu,J.R. (2000). MAP kinases in fungal pathogens. *Fungal Genet. Biol.* 31, 137-152.
16. Zeigler,R.S., Leong,S.A., and Teng,P.S. (1994). Rice blast disease. (Kew, UK: CAB International Mycology Institute).

Appendix

Secretome display

Objective

A method was developed based on the yeast surface display system to identify secreted protein from eukaryotes. In the yeast display system, aga2p, a subunit of agglutinin receptor, is normally expressed as an N-terminal fusion (Colby et al., 2004). Aga1p is a GPI-anchored surface protein in budding yeast, to which Aga2p is covalently bound by two disulfide bonds. We found that aga2p can be expressed, and displayed as a C-terminal fusion on the yeast surface (Figure A.1A). Fluorescence Activated Cell Sorting (FACS) and Magnetic Activated Cell Sorting (MACS) strategies were used for the selection and enrichment of yeast cells expressing cDNAs encoding a signal peptide which were tagged at C-terminus with Aga2p. This method provides a potential strategy for the enrichment of secreted proteins from a complex cDNA library.

Materials and Methods

The secretome display vector is based on pYD1 yeast display vector (Invitrogen). Using the pYD1 as a template, PCR was done using primers (promoter 5'-GAAGAACC GCGGGAAAATGTATGAAGTAGAATTGC-3' and aga2 mature protein sequence 5'-ATCATCCCGCGGCAGGAACTGACAACTATATGC-3') to eliminate the signal peptide sequence from Aga2 coding region. The PCR product was digested with SacII restriction enzyme (Promega), self-ligated, and transformed into *E. coli*, generating pSECDIS. The functionality of markers such as the ampicillin cassette and the Trp yeast nutritional marker were functionally tested. pSECDIS777 was constructed by inserting

FLAG and spacer region of pSPDK777 (a kind gift from Dr. Dinesh-Kumar at Yale University) in to the SacII site of pSECDIS depicted in Figure A.1B. The FLAG and spacer region were PCR amplified using the primers 777F (GACGAACCGCGG **CCCGGG**GAGCTCAAGCTTCTCGAGGAT) and 777R (GAACCACCGCGGTACCG AGCTCGAATTCGCGTC). A GATEWAY-compatible vector was constructed by inserting the GATE- WAY conversion cassette into the SmaI site (the sequence in bold character in the 777F primer) of pSECDIS777 in frame; we named the resulting vector pSECDIS777RC. MPG1 coding sequence with and without signal peptide was cloned into pSECDIS777 (A: pSECDIS777-MPG+SIG, B: pSECDIS777-MPG1-SIG). Cells were incubated with anti-V5 antibody (Invitrogen) and anti-FLAG antibody (Sigma) after blocking with Bovine serum albumin (BSA) (1mg/ml) in phosphate buffered saline (PBS). To minimize the false positive selection, two colors fluorescence activated cell sorting (FACS) was done using secondary antibodies conjugated to R-phycoerythrin (PE) for FLAG tag and fluorescein isothicyanate (FITC) for V5 tag. Cells were extensively washed with PBS. FACS was performed using DakoCytomation MoFlo cell sorter. A cDNA library was prepared from appressoria of the rice blast fungus, *Magnaporthe grisea* 70-15 strain using random hexanucleotide priming. The plasmid library in *E. coli* DH10B was prepared and transformed into *S. cerevisiae* EBY1000. The yeast culture was grown in synthetic media without Trp/Ura with glucose as a carbon source (SC-GLU-(Trp/Ura)) to log phase, and collected by centrifugation at room temperature. The cell pellet was washed two times with sterile distilled water. Induction of gene expression was done in SC+GAL-(Trp/Ura) at 20C for 12-36 hrs. Magnetic bead enrichment was done as follows: Induced cells were treated with anti-V5 antibody followed by treatment with

DSB-X-biotin-conjugated secondary antibody. DSB-X-biotin conjugation, unlike biotin-streptavidin binding, provides a reversible binding to streptavidin. Cells were washed with PBS and binding to streptavidin magnetic beads were performed at room temperature with gentle shaking. Unbound cells were washed by extensive washing with PBS, and bound cells were released by adding biotin solution (10 mM final). The released cells were plated on SC-GLU-(Trp/Ura) medium.

Results and discussion

To develop a yeast display system to detect secreted and membrane-associated proteins from *M. grisea*, we first constructed a vector expressing Aga2p as a C-terminal fusion, pSECDIS777 (without signal peptide). MPG1 coding sequence with or without the signal peptide was cloned into the vector. Two selection strategies were tested. Staining of positive (pYD1) and negative (pSECDIS) controls with anti-V5 antibody and secondary staining with FITC-conjugated antibody resulted in a clear distinction (Figure A.2A). Fluorescence from cells harboring the MPG1 constructs with a signal peptide was readily observed under the fluorescence microscope. No fluorescence was detected from cells in which the signal peptide has been deleted from MPG1. FACS analysis also showed that MPG1+Sig construct produced a distinct population of yeast cells compared to cells transformed with MPG1-Sig construct, clearly indicating that MPG1::Aga2p fusion protein is expressed and correctly displayed on the yeast cell wall (Figure A.2B).

Encouraged by the results obtained, we constructed an appressoria cDNA library in pSECDIS777. Yeast cells transformed with the cDNA library was subjected to FACS analysis. For the first round of FACS selection, no distinct population was observed.

However, second round selection of FACS using cells obtained from the terminal tip of first round selection produced a distinct population (Figure A.3A). The cells were collected, and induced again to confirm the expression of fusion protein by immunostaining (anti-V5). Selected cells were stained brightly with anti-V5 antibody. In contrast, residual cells exhibited poor staining. We explored enrichment based on affinity separation using magnetic bead. Cells were incubated with anti-V5 antibody followed by biotin-conjugated antibody. Finally, streptavidin-coated magnetic beads were added to antibody-treated cells. As shown in Figure A.3B, cells recovered from magnetic beads following induction resulted in significantly more yeast colonies compared to uninduced cells, strongly suggesting the utility of the enrichment strategy for recovering yeast cells displaying fusion proteins. Thus, our data indicates that *aga2p* can be expressed and anchored on the yeast cell wall as a C-terminal fusion, and can be a powerful tool for genome-wide screens to identify secreted proteins. The use of yeast as a host for protein expression may be most applicable to the study of eukaryotic proteins. However, caution should be exercised since eukaryotic protein secretion is under strict quality control such as unfolded protein response (UPR), and may not efficiently produce foreign proteins. In addition, signal peptides from heterologous genes may not be recognized by yeast secretion system. Thus, UPR and properties of signal peptide might be technical issues to be resolved (Mattanovich et al., 2004).

References

1. Colby,D.W., Kellogg,B.A., Graff,C.P., Yeung,Y.A., Swers,J.S., and Wittrup,K.D. (2004). Engineering antibody affinity by yeast surface display. *Methods Enzymol.* 388, 348-358.
2. Mattanovich,D., Gasser,B., Hohenblum,H., and Sauer,M. (2004). Stress in recombinant protein producing yeasts. *J Biotechnol.* 113, 121-135.

Figure legends

Figure A.1 Expression of Aga2p as a C-terminal fusion.

(A) Aga2p coding sequence without start codon and signal peptide is translationally fused to a foreign cDNA sequence. If the foreign cDNA provides a start codon and a signal peptide, cDNA::Aga2p is expressed and exported to the cell surface. Aga1p is a GPI-anchor protein to which Aga2p is covalently linked by two disulfide bonds (indicated by red bars). (B) Configuration of pSECDIS777.

Figure A.2 Expression of MPG1 test constructs and FACS analysis.

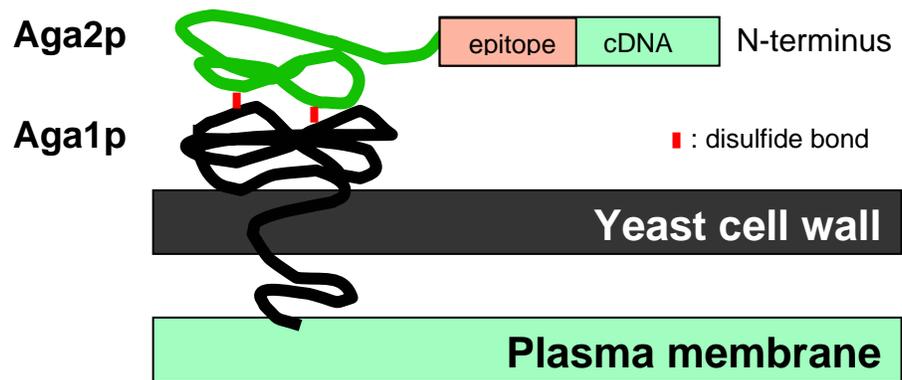
(A) MPG1 coding sequences with or without a signal peptide (MPG1+Sig and MPG1-Sig, respectively) were cloned into pSECDIS777. Induced cells were stained with FITC-conjugated antibody. In the positive control (pYD1) and MPG1+Sig, bright staining is observed. In the negative control and MPG1-Sig, background fluorescence is observed, indicating that MPG1 with a signal peptide fusion protein is expressed and displayed on the cell wall. Upper panels are picture of cells observed under bright field microscope, and lower panels are cells observed with FITC. (B) FACS analysis after staining the MPG1+Sig and MPG1-Sig transformed cells with R-PE. Formation of a distinct population is shown in MPG1+Sig (left panel) compared to MPG1-Sig (right panel).

Figure A.3 Enrichment of cells expressing proteins with signal peptide by FACS and MACS.

A complex cDNA library was generated from appressoria produced by *M. grisea*. Yeast cells were transformed with the library, and expression was induced. (A) No apparent distinct population is visible in the first round of FACS. However, in the second round, a

distinct population is clearly evident. (B) Cells obtained from the second round of FACS selection was induced, and compared with the residual cells. In selected cells, fluorescence is evident after staining with anti-V5 antibody followed by FITC-conjugated antibody staining. However, background level of staining is observed in residual cells. (C) MACS strategy was tested for the enrichment of cells displaying cDNA::Aga2p fusion proteins. Over 100-fold more colonies are observed in the induced cells (right side) compared to uninduced cells (left side),

(A)



(B)

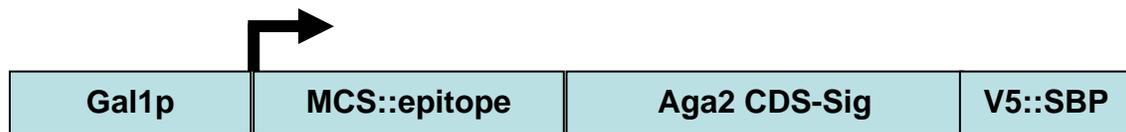
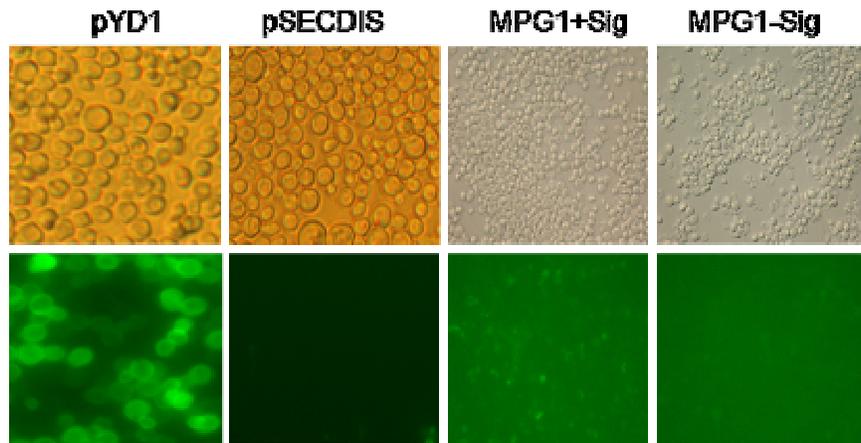


Figure A.1 Expression of Aga2p as a C-terminal fusion

(A)



(B)

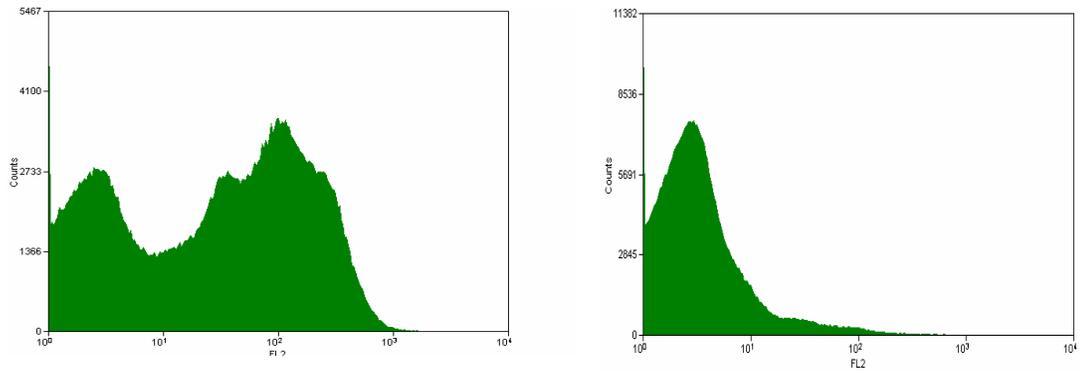
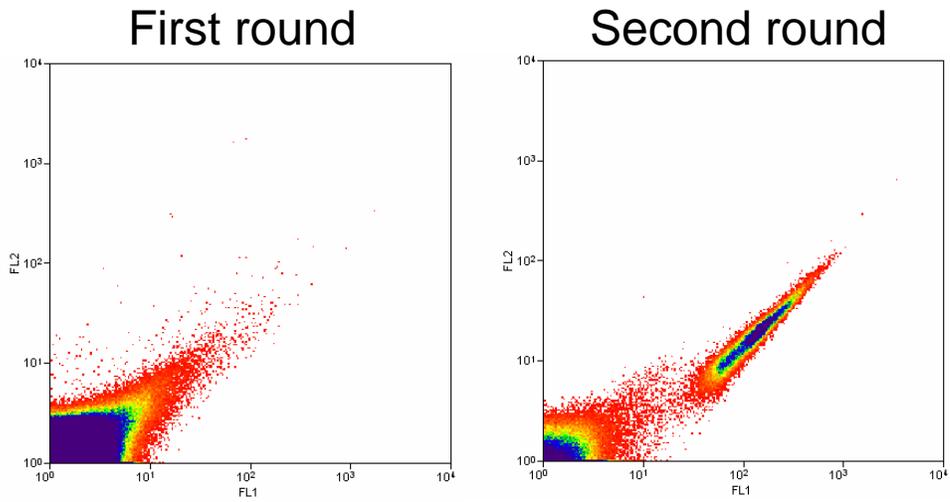
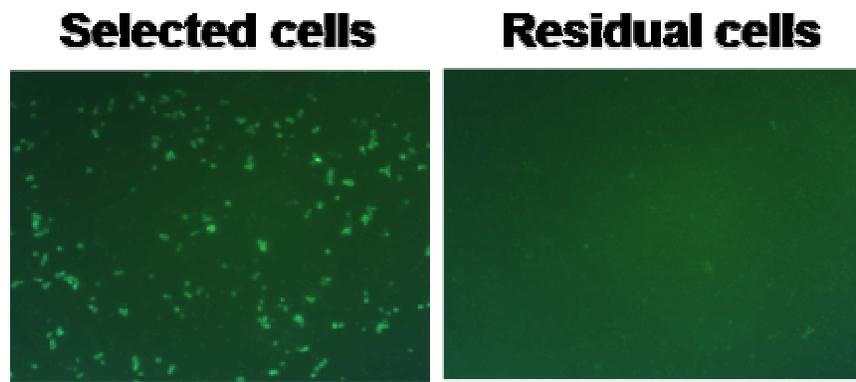


Figure A.2 Expression of MPG1 test constructs and FACS analysis

(A)



(B)



(C)

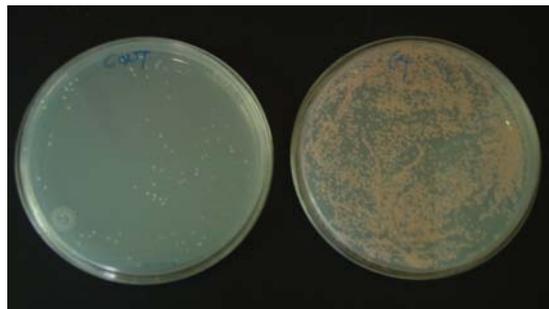


Figure A.3 Enrichment of cells expressing proteins with signal peptide by FACS and MACS