MOLECULAR CHARACTERIZATION OF THE MGENP AVIRULENCE 
GENE FAMILY IN THE BLAST FUNGUS MAGNAPORTHE GRISEA

A Thesis in
Plant Pathology
by
Chang Hyun Khang

© 2005 Chang Hyun Khang

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2005
The thesis of Chang Hyun Khang was reviewed and approved* by the following:

Seogchan Kang
Associate Professor of Plant Pathology
Thesis Advisor
Chair of Committee

David M. Geiser
Associate Professor of Plant Pathology

Gary W. Moorman
Professor of Plant Pathology

Gretchen A. Kuldau
Assistant Professor of Plant Pathology

Timothy W. McNellis
Assistant Professor of Plant Pathology

Surinder Chopra
Assistant Professor of Maize Genetics

Leonard J. Francl
Professor of Plant Pathology
Head of the Department of Plant Pathology

*Signatures are on file in the Graduate School
ABSTRACT

The haploid heterothallic ascomycete *Magnaporthe grisea* (Herbert) Barr causes blast diseases on many gramineous species. This fungal species can be differentiated into several genetically distinct host specific groups. Rice pathogenic *M. grisea* strains cause rice blast disease that limits rice production worldwide. One desirable strategy for controlling the disease is to use blast-resistant rice cultivars. However, resistance has not been durable because of the frequent appearance of new races in the field with the ability to infect previously resistant cultivars. Because the interactions between *M. grisea* and rice follow the gene-for-gene model, the mutations of avirulence (*AVR*) genes can lead to the generation of new races. Thus, understanding the molecular mechanisms underpinning the evolution and variation of *AVR* genes is critical to understanding how the pathogen overcomes host resistance.

The *AVR* gene *MgMEP1* (previously called *AVR-Pita*), which was cloned from a rice pathogenic field isolate of *M. grisea* (O-137), prevents the fungus from infecting rice cultivars containing the corresponding *R* gene Pi-ta. A survey of more than 300 *M. grisea* strains isolated from diverse hosts identified sequences homologous to the *MgMEP1* gene with varying degrees of sequence similarity, suggesting that the *MgMEP1* gene is a member of a gene family. Two other family members, termed *MgMEP2* (functional as an *AVR* gene) and *MgMEP3* (nonfunctional as an *AVR* gene), were cloned based on their homology to the O-137 *MgMEP1* gene. The *MgMEP1* and *MgMEP2* genes were present widely even among strains that are not pathogenic to rice and genetically distant from the rice pathogens. Among these strains, substantial variation was observed with respect to the gene copy number (ranging from zero to three) and gene-linked restriction fragment length polymorphisms (RFLPs). The *MgMEP3* gene was present only among rice pathogens and those that are genetically close to rice pathogens (forming a new phylogenetic species termed *M. oryzae*) and did not exhibit much variation in its copy number and RFLPs. This suggests that the *MgMEP3* gene was the newest member of the gene family generated only after the divergence of *M. oryzae* from *M. grisea* and is genetically more stable than the *MgMEP1* and *MgMEP2* genes. Sequence analysis of
avirulence and virulence alleles of the *MgMEP* gene family members showed that most of them are likely to encode a functional metallopeptidase, suggesting that the evolution of the *MgMEP* gene family involves modifications of a protein to preserve enzymatic function while altering its recognition specificity as an avirulence factor.

To determine the molecular mechanisms underpinning the genetic variability of the gene family, the genomic context of its members was analyzed in eight different *M. grisea* strains isolated from diverse hosts. The *MgMEP1/MgMEP2* (probably allelic) genes were located near a telomere and flanked by diverse DNA repeats. This suggests that frequent deletion/amplification of the *MgMEP1/MgMEP2* genes have resulted from their telomeric location and recombination mediated by the DNA repeats flanking them. The relative stability of the *MgMEP3* gene was probably due to its internal location and absence of DNA repeats at flanking regions.

Rapid progress in fungal genome sequencing presents many new opportunities for studying fungal biology through a systematic mutagenesis of the genes identified through sequencing. However, the lack of efficient tools for targeted gene replacement is a limiting factor for fungal functional genomics, as it often necessitates the screening of a large number of transformants to identify the desired mutant. An efficient method of gene replacement was developed, and factors affecting the efficiency of this method were evaluated using two plant pathogenic fungi, *Magnaporthe grisea* and *Fusarium oxysporum*. This method is based on *Agrobacterium tumefaciens*-mediated transformation with a mutant allele of the target gene flanked by the herpes simplex virus thymidine kinase (*HSVtk*) gene as a conditional negative selection marker against ectopic transformants. The *HSVtk* gene product converts 5-fluoro-2’-deoxyuridine to a compound toxic to diverse fungi. Because ectopic transformants express *HSVtk*, while gene replacement mutants lack *HSVtk*, growing transformants on a medium amended with 5-fluoro-2’-deoxyuridine facilitates the identification of targeted mutants by counter-selecting against ectopic transformants. In addition to *M. grisea* and *F. oxysporum*, the method and associated vectors are likely to be applicable to manipulating genes in a broad spectrum of fungi, thus potentially serving as an efficient, universal functional
genomic tool for harnessing the growing body of fungal genome sequence data to study fungal biology.
TABLE OF CONTENTS

List of Figures........................................................................................................ x

List of Tables ........................................................................................................ xi

Acknowledgements …........................................................................................... xii

Chapter 1: General Introduction ........................................................................ 1
  1.1 Introduction.................................................................................................. 1
  1.2 Literature Review......................................................................................... 4
    1.2.1 Pathogen avirulence (AVR) genes and effector genes.................... 4
    1.2.2 Structure and properties of fungal AVR genes............................... 6
    1.2.3 Models of AVR gene function.......................................................... 7
    1.2.4 Role of AVR genes for pathogen fitness......................................... 9
    1.2.5 Evolution of AVR genes................................................................... 11
      1.2.5.1 Loss of AVR gene function ................................................... 12
      1.2.5.2 Divergence (Alternation) of AVR gene function..................... 13
  1.3 Conclusion ................................................................................................. 16
  1.4 Outline of the Thesis.................................................................................. 16

Chapter 2: Characterization of the MgMEP Avirulence Gene Family
  among Isolates of Magnaporthe grisea from Diverse Hosts ...................... 18
  2.1 Introduction.................................................................................................. 18
  2.2 Materials and Methods................................................................................ 20
    2.2.1 Strains and growth conditions.......................................................... 20
    2.2.2 Cloning of the MgMEP homologs..................................................... 22
      2.2.2.1 Cloning by genomic DNA library screening............................. 22
      2.2.2.2 Cloning by PCR .................................................................... 22
      2.2.2.3 Cloning in silico .................................................................... 24
Chapter 3: Comparative Analysis of Genomic Organization of the \textit{MgMEP} Gene Family among Isolates of \textit{Magnaporthe grisea} from Diverse Hosts

3.1 Introduction ........................................................................................................ 41

3.2 Materials and Methods .................................................................................. 43

3.2.1 Strains and growth conditions .................................................................... 43

3.2.2 Cloning by genomic library screening ....................................................... 44

3.2.3 Inverse PCR ................................................................................................ 44

3.2.4 DNA sequencing and computer analysis .................................................... 45

3.2.5 Database searching ...................................................................................... 45

3.3 Results ............................................................................................................. 46

3.3.1 Some members of the \textit{MgMEP} gene family are closely linked to the telomere .............................................................................................................. 46
3.3.2 Genomic location of the \textit{MgMEP1} and \textit{MgMEP3} genes in \textit{M. grisea} strain 70-15 ................................................................. 49
3.3.3 Identification and characterization of two novel gypsy-type retrotransposons that are closely associated with \textit{MgMEP} genes ................................................................. 50
3.3.4 Syntenic relationship in the regions containing the \textit{MgMEP1} and \textit{MgMEP2} genes ............................................................................................. 55
3.3.5 Syntenic relationship in the regions containing the \textit{MgMEP3} gene in G-17, G-22, and 70-15 ................................................................. 55
3.3.6 Syntenic relationship between the regions of the \textit{MgMEP1}/\textit{MgMEP2} genes and \textit{MgMEP3} genes ............................................................................. 56
3.4 Discussion ............................................................................................ 57

Chapter 4: Dual Selection Based, Targeted Gene Replacement Tool for \textit{Magnaporthe grisea} and \textit{Fusarium oxysporum} ........................................ 62
4.1 Introduction ....................................................................................... 62
4.2 Materials and Methods ..................................................................... 64
  4.2.1 Strains, media, and ATMT-DS ..................................................... 64
  4.2.2 Vector construction .................................................................. 65
4.3 Results ............................................................................................. 68
  4.3.1 Herpes simplex virus thymidine kinase (HSVtk) functions as a negative selection marker in diverse fungi .............................................. 68
  4.3.2 Targeted gene replacement in \textit{F. oxysporum} and \textit{M. grisea} genes using ATMT-DS ................................................................. 71
  4.3.3 Stability of \textit{HSVtk} depends on fungal strains and its location in the T-DNA ................................................................. 74
  4.3.4 Construction of improved targeted gene replacement vectors...... 75
4.4 Discussion ....................................................................................... 76
4.5 Acknowledgements ........................................................................... 80
Chapter 5: General Discussion .......................................................................... 81

Appendix .............................................................................................................. 84

Abstract ............................................................................................................. 84
1. Introduction ................................................................................................... 85
2. Materials ....................................................................................................... 87
2.1. Fungal and Bacterial Strains .............................................................. 87
2.2. Culture Media .................................................................................... 87
2.3. Antibiotics and Other Chemical Agents ........................................... 89
2.4. Binary Vectors for Fungal Transformation ....................................... 89
2.5. Other Solutions and Supplies............................................................. 90
3. Methods........................................................................................................ 91
3.1. Transformation of \( A. tumefaciens \) with Binary Vector............... 91
3.1.1. Preparation of Competent Cells .............................................. 91
3.1.2. Transformation ....................................................................... 92
3.2. Transformation of \( M. grisea \) and \( F. oxysporum \) ................ 93
3.2.1. Preparation of \( M. grisea \) Spores.............................................. 93
3.2.2. Preparation of \( F. oxysporum \) Spores ..................................... 93
3.2.3. Transformation................................................................. 94
3.2.4. Isolation and Purification of Transformants ......................... 94
3.2.5. Molecular Analysis of Transformants .................................... 95
3.2.5.1. Southern Analysis ........................................................... 95
3.2.5.2. Identification of Gene Knock-Out Mutants by PCR ...... 96
3.2.5.3. Thermal Asymmetric Inter-Laced PCR........................... 97
3.3. Targeted Gene Knock-Out via ATMT-DS ........................................ 99
3.3.1 Transformation......................................................................... 99
4. Notes ............................................................................................................ 99
Acknowledgements ......................................................................................... 106

Bibliography ...................................................................................................... 107
LIST OF FIGURES

Figure 1. Distribution pattern of the MgMEP gene family in Magnaporthe grisea isolates from diverse hosts ................................................................. 29

Figure 2. Comparison of avirulent and virulent MgMEP genes .................. 33

Figure 3. Gene genealogy of 23 MgMEP genes ................................................ 36

Figure 4. Comparison of the genomic organizations of the MgMEP loci in different M. grisea isolates ................................................................. 48

Figure 5. Restriction map and organization of the G-213 MgMEP2 locus and Southern blot analysis of genomic DNA from 27 Magnaporthe grisea field strains ..................................................................................... 54

Figure 6. Schematic diagram of the Fusarium oxysporum FoSNF1 and Magnaporthe grisea MHP1 loci ................................................................. 66

Figure 7. Schematic diagram of ATMT-DS .................................................. 69

Figure 8. Schematic diagrams of the T-DNA regions in the binary vectors ........ 69

Figure 9. Growth of Magnaporthe grisea, Fusarium oxysporum, Aspergillus fumigatus, and Botrytis cinerea in the presence of 5-fluoro-2'-deoxyuridine (F2dU) or ganciclovir ................................................................. 70

Figure 10. Southern analysis of Magnaporthe grisea ATMT-DS transformants .... 73

Figure 11. Stability of neo and HSVtk located at either LB or RB of the T-DNA ...... 75

Figure 12. Schematic diagrams of the T-DNA region in five binary vectors for fungal transformation ................................................................. 90

Figure 13. PCR-based screening of gene KO mutants ...................................... 97

Figure 14. Primers for TAIL-PCR .............................................................. 98
LIST OF TABLES

Table 1. *Magnaporthe grisea* field isolates used in this study.............................21

Table 2. Primers used in this study .................................................................23

Table 3. Pairwise sequence comparison between *MgMEP* genes.......................30

Table 4. Efficiency of targeted gene replacement via ATMT-DS in *M. grisea* and *F. oxysporum* ..................................................................................72

Table 5. Leakiness of direct negative selection during the regeneration of transformants .........................................................................................74

Table 6. Preparation of stock solutions and transformation media .......................88
ACKNOWLEDGMENTS

I would like to express my gratitude to those who have inspired me during my education. Most of all, I am greatly indebted to my advisor, Dr. Seogchan Kang for his guidance and encouragement throughout my study. Especially, his knowledge and enthusiasm for exploring new ideas inspired me in every moment of my life here. I am grateful to him for being a professional role model. I also extend my sincere thanks to the other members of my committee, Drs. Gary Moorman, David Geiser, Gretchen Kuldau, Timothy McNellis, and Surinder Chopra for their constructive advice and willingness to share their insight and wisdom all along the way. I would also like to thank Dr. Fred Gildow for being my teaching mentor and sharing his teaching experience with me. My special thanks goes out to Dr. Yong-Hwan Lee, without whose encouragement and inspiration I would not have experienced the joy of doing science. I would like to thank Dr. Barbara Valent for sharing unpublished data and thoughtful insights into the AVR-Pita project.

I would like to acknowledge the department head, Dr. Leonard Francl and all the personnel in the Plant Pathology Department for their support during my education. I am especially appreciative of Roxanne Lease for her assistance in my greenhouse experiment. I also extend my thanks to my fellow graduate students and postdocs for their encouragements and friendships. I must also thank my long-time friends and colleagues, Dr. Sook-Young Park and Dr. Hee-Sool Rho for their assistance and helpful suggestions.

Finally, I would like to thank my parents and sisters for their prayer and support through my life. I also want to thank my fiancé and best friend, Ah-Young Cho for her prayer and encouragement. Thanks and recognition are also extended to my spiritual teachers and friends for sharing the gospel and guiding me to find the Truth in Christ.

This research was supported by grants from USDA-NRI, Crop Functional Genomics Center of the 21st Century Frontier Research Program, and the Korea Science and Engineering Foundation.
Chapter 1

GENERAL INTRODUCTION

1.1 Introduction

Plant diseases are of major importance to humans because they can damage plants and plant products that provide food, clothing, housing, and hospitable environment for sustaining human societies (AGRIOS 1997). The socio-economic consequences of plant diseases have been well illustrated by many diseases in the past, including the Irish potato famine in 1845 and the great Bengal famine in 1943 (AGRIOS 1997). In today’s world, 852 million people suffer from chronic hunger, according to the Food and Agriculture Organization (FAO) of the United Nations (FAO 2004). Despite numerous control methods available, at least one third of global crop production is destroyed by a multitude of pathogens and pests during pre- and post-harvest stages (AGRIOS 1997). As world population continues to increase and agroecosystems become less productive, protection of plants from diseases, substantially leading to improved agricultural production, will be one of the most important necessities for providing an overpopulated world with food as well as other plant products. Pesticides have provided effective protection, but their continuing and increasing uses have posed a threat to human health and the environment. In addition, pesticide application in developing countries is often not economically feasible (LAPPE et al. 1998). For these reasons, much effort has been made to develop environmentally sound and economically practical control strategies by understanding and utilizing the plants’ own defense mechanisms.

Plants protect themselves from potential pests/pathogens using preformed and induced defense responses (AGRIOS 1997). Preformed defense responses comprise constitutive structural or chemical attributes of plants, regardless of the presence or absence of pathogen attacks, and include physical barriers, presence of antimicrobial compounds or lack of nutrients or targets for virulence factors required for potential
pathogens. Induced defense responses, on the other hand, are activated upon the recognition by the plant cells of the invading pathogens. This recognition event triggers a signal transduction cascade, leading to defense responses that include production of phytoalexins (Dixon 1986), hydrolytic enzymes (Boller 1987), pathogenesis-related proteins (Sticher et al. 1997), protease inhibitors (Ryan 1990), reactive oxygen species (Mehdy 1994), the deposition of lignin (Friend 1976) and callose (Kauss 1987) into the plant cell wall, and a rapid and localized cell death at the infection site (hypersensitive response, HR) (Hammond-Kosack and Jones 1996; Stakman 1915). These local responses often lead to a long lasting systemic response (systemic acquired resistance, SAR) that provides resistance to a wide range of pathogens (Dong 2001).

The most common form of disease resistance is nonhost resistance (broad resistance), in which an entire plant species is resistant to all members of a given pathogen species (Heath 1991, 1996, 2000). Nonhost resistance seems to be pathogen-nonspecific and involves preformed and induced defense responses. In general, complex genetic controls appear to be involved in nonhost resistance. On the other hand, host resistance (cultivar-specific resistance) is expressed by certain plant genotypes of otherwise susceptible host species to a particular pathogen (Heath 1991). Host resistance is often conditioned by single resistance (R) genes of host plants with their products directly or indirectly recognizing the specific elicitors produced by corresponding avirulence (AVR) genes of pathogens (Hammond-Kosack and Jones 1996). This specific recognition results in induction of defense responses including HR. This genetic relationship of host and pathogen was first formulated by H.H. Flor in his gene-for-gene hypothesis, based on his studies on the genetic inheritance of pathogenicity of the flax rust fungus Melampsora lini on different flax cultivars (Flor 1946, 1947, 1955). His work demonstrated that a flax cultivar is resistant to a particular race of M. lini only if the cultivar carries a dominant R gene that corresponds to a dominant AVR gene of the pathogen. Since then, gene-for-gene relationships have been demonstrated in many plant-pathogen systems, including viruses, bacteria, nematodes, and insects (Agrios 1997).

Because single R genes can induce effective defense responses that lead to complete resistance against particular pathogens, R genes have been employed in resistance
breeding programs for decades (Pink 2002). R gene-mediated resistance is a desirable method to control diseases for several reasons. Single R genes can be easily introduced into many crops using transgenic techniques. When induced as a result of specific interaction with pathogen AVR genes, concerted defense responses efficiently restrict pathogen growth at the site of infection. This method is convenient and economical for farmers because no investments in machinery, labor, and special training are required.

However, R gene-mediated resistance is often defeated by a new race of pathogen. As a resistant cultivar containing an R gene is widely cultivated, selection pressure increases against the pathogen populations containing the corresponding AVR gene. The selection pressure results in the elimination of those members of the population containing the AVR gene product, leaving those members of the population that lack the host-recognizable product to increase in the population. As a result, the virulent pathogen race increases that can overcome the resistance. Although there are some cases demonstrating that single R gene-mediated resistance has remained effective over several decades (Leach et al. 2001), such resistance is ultimately defeated by mutations of the pathogen AVR genes. To improve the durability of R gene-mediated resistance, as an alternative to deployment of a single R gene, several strategies have been proposed, which include gene pyramiding (several R genes into a single cultivar) (Crute and Pink 1996; Mundt 1990; Pederson and Leath 1988), multilines (mixtures of near-isogenic lines with different R genes) (Browning and Frey 1969), and cultivar mixtures (mixtures of heterogeneous cultivars with different R genes, but with similar agronomic traits to be grown together) (Wolfe and McDermott 1994). These strategies are based on the idea that multiple R genes, either in a single cultivar or in mixtures of cultivars, require the pathogen population to accumulate mutations in multiple corresponding AVR genes to overcome the resistance. This is unlikely to occur if the mutations impose a cumulative fitness penalty on the pathogen itself (discussed below; Pink 2002)). These approaches will become more practical as more R genes are cloned and characterized, and plant transformation technology is improved. A major concern in using gene pyramiding is that there would be strong unidirectional selection against the matching AVR genes that lead to evolution of a virulent pathotype (Pink 2002). In using cultivar
mixtures / multilines, there is a concern that the continuous deployment would select for complex races (Groth 1976). Thus, monitoring the pathogen population for the frequency of the mutated AVR genes will be required for the effective deployment of such strategies for disease control.

Increased knowledge of AVR and R genes, their products, and their molecular interactions provides insight into plant defense mechanisms, and opportunities for developing novel and improved strategies for the control of plant diseases (Laugé and De Wit 1998). Durability of the R gene-mediated resistance can be achieved by better understanding the mechanisms underpinning AVR gene evolution and variation. In this review on fungal AVR genes, structural properties of AVR genes, models of AVR gene function, pathogenicity function of AVR genes, and evolution of AVR genes will be summarized.

1.2 Literature Review

1.2.1 Pathogen avirulence (AVR) genes and effector genes

It is intuitive that a pathogen carries genes for pathogenicity that enable it to parasitize host plants. However, the idea that pathogens carry genes for avirulence, which prevent the pathogens from infecting certain hosts, is less so. The mechanisms of host resistance, pathogen variation, and coevolution between host and pathogen can be clearly explained by the pathogen avirulence (AVR) genes, which were first proposed in the gene-for-gene hypothesis proposed by H. H. Flor (Flor 1946; Flor 1947). His studies demonstrated that on varieties of flax containing a gene for resistance (R gene) to the avirulent parent race (morphologically indistinguishable pathogen variant within a pathogen species), the F2 progeny of the fungus segregated in a monofactorial ratio of 1 virulent : 3 avirulent. Thus, pathogen virulence was recessive and pathogen avirulence was dominant. On the varieties that have more than one R gene to the avirulent parent
race, the F2 progeny of the fungus segregated in ratios according to the corresponding numbers of \( R \) genes in host varieties. This suggests that for each gene conditioning resistance in the host, there is a corresponding gene conditioning virulence in the pathogen (gene-for-gene relationship). In his genetic studies, the recessive virulence gene was always emphasized as the active determinant that conditioned cultivar-specific virulence rather than the dominant \( AVR \) genes, probably because of the dilemma posed by the concept that pathogen avirulence was determined by its own dominant gene (GABRIEL 1999). Flor’s genetic studies showing that \( AVR \) genes were genetically dominant led to the subsequent molecular cloning of the first \( AVR \) gene from a bacterial pathogen \textit{Pseudomonas syringae} (STASKAWICZ et al. 1984). Since then, over 40 \( AVR \) genes have been cloned from gram-negative bacterial pathogens of the genera \textit{Pseudomonas}, \textit{Xanthomonas}, \textit{Erwinia}, and \textit{Ralstonia} (for review see (LEACH and WHITE 1996; VAN’T SLOT and KNOGGE 2002)). About 18 \( AVR \) genes have been cloned from diverse fungi, including \textit{Magnaporthe grisea}, \textit{Cladosporium fulvum}, \textit{Rhynchosporium secalis}, and \textit{Melampsora lini}, and Oomycetes, including \textit{Phytophthora infestans} and \textit{P. parasitica} (for review see (LAUGÈ and DE WIT 1998; VAN’T SLOT and KNOGGE 2002)). The cloned \( AVR \) genes determine cultivar specificity in that \( AVR \) genes, when transferred into a virulent pathogen race, render the recipient race avirulent to the host cultivars that carry corresponding \( R \) genes. This fits the original gene-for-gene hypothesis that predicted the dominance of the \( AVR \) gene. Disruption of an \( AVR \) gene makes an avirulent pathogen race virulent to the previously resistant cultivar, demonstrating that cultivar specific virulence is conditioned by the mere absence of the \( AVR \) gene. This contrasts with Flor’s prediction that a recessive virulence gene conditions cultivar specific virulence. Thus, the historical term ‘virulence genes’ referring to hypothetical alternate allele of cultivar specific \( AVR \) genes appears to be a misnomer since they do not condition the cultivar-specific virulence (GABRIEL 1999).

\( AVR \) genes are often distinguished from pathogenicity genes. However, many \( AVR \) genes have been demonstrated to be required for full virulence on the host lacking the \( R \) gene (for review see (VAN’T SLOT and KNOGGE 2002)). This suggests that \( AVR \) genes have dual roles as being pathogenicity factors as well as avirulence factors. When
pathogen genes trigger resistance responses on the host carrying R genes, they are termed AVR genes. At the same time, they are required for virulence on the host lacking R genes. To overcome the confusing terminology for the same gene, the new term ‘effector’ gene has been adapted to designate pathogen genes that are involved in the plant-pathogen interaction, regardless of their function as avirulence or pathogenicity factors (Van Dijk et al. 1999). Because many pathogen genes have been first characterized as avirulence determinants, many of them are still called AVR genes. The term ‘AVR gene’ will be used here rather than ‘effector gene’.

1.2.2 Structure and properties of fungal AVR genes

To date 18 fungal AVR genes have been cloned and characterized. When their DNA and protein sequences are compared, there seems to be little similarity. However, some of them share common features such as cysteine-rich regions and signal peptides (Van't Slot and Knogge 2002). Cysteine pairs can form intramolecular disulfide bonds that stabilize the proteins. Signal peptides are short discrete stretches of amino acid sequence, which are part of a protein and direct proteins to their proper cellular and extracellular locations. The AVR genes Avr2, Avr4, and Avr9 from the tomato leaf mold Cladosporium fulvum (Joosten and de Wit 1999) and NIP1 (Necrosis-Inducing Peptide) from the barley scald fungus Rhynchosporium secalis (Rohe et al. 1995) encode low-molecular weight proteins. These proteins contain an even number of cysteine residues for disulfide bonds and a signal peptide for extracellular targeting. It has been shown that the substitution of one single cysteine residue of Avr4 and Avr9 significantly compromised the HR-inducing activity of the proteins (Joosten et al. 1997; Kooman-Gersmann et al. 1997). The cysteine residues of Avr9 have been also shown to be involved in disulfide bonds (Van den Hooven et al. 2001). Thus, cysteine residues appear to be important for the stable protein structure and for the full avirulence activity. All of these AVR proteins are localized in the apoplastic space, indicating the role of their signal peptide sequences for secretion. Signal peptides are also found in other AVR gene products including AVR-Pita and PWL (Pathogenicity toward Weeping Lovegrass) genes from M. grisea (Kang
et al. 1995; ORBACH et al. 2000; SWEIGARD et al. 1995), and AvrL567 from the flax rust fungus *Melampsora lini* (DODDS et al. 2004).

Most of the fungal *AVR* genes encode proteins that have no similarity to the characterized proteins in the databases. Only two *AVR* genes, *AVR-Pita* (ORBACH et al. 2000) and *ACE1* (for Avirulence Conferring Enzyme1) (BÖHNERT et al. 2004) from *M. grisea*, encode proteins that show similarity to zinc metalloprotease and polyketide synthase, respectively. These enzymatic activities appear to be responsible for their avirulence activity (discussed below).

1.2.3 Models of *AVR* gene function

Several models have been proposed for the biochemical basis of *AVR* gene function. In the simplest elicitor-receptor model, the *AVR* gene product is an elicitor or ligand that is directly recognized by the receptor produced by the corresponding *R* gene, which is located outside or inside the host cell. This specific recognition activates a signal transduction pathway that leads to resistance responses, often involving HR (GABRIEL and ROLFE 1990; JONES 1997). Only a few *AVR* gene products have been demonstrated to interact directly with their corresponding *R* gene products (DESLANDES et al. 2003; JIA et al. 2000; TANG et al. 1996). The lack of evidence for direct *AVR*-R interactions led to the ‘guard model’, in that the *AVR* gene product targets a third protein of the host, which is likely to be the host target for pathogen virulence factor. The interaction of the AVR with such host proteins is then recognized by the R gene product (DANGL and JONES 2001; VAN DER BIEZEN and JONES 1998). Supporting evidence for this model has been accumulating by the identification of such targets in several gene-for-gene interactions (KIM et al. 2002; SCHNEIDER 2002). In the second model, the *AVR* gene product is an enzyme that directs the synthesis of an elicitor-active compound(s). The *avrD* from the bacterial pathogen *P. syringae* pv. *tomato* has been demonstrated to direct the extracellular accumulation of low molecular weight acyl glycosides that are sufficient to trigger the HR (KEEN et al. 1990; MIDLAND et al. 1993). Another example is the *ACE1* (for Avirulence Conferring Enzyme1; formerly *AVR-Irat7*), an *AVR* gene cloned from *M.*
oryzae that conditions avirulence on a rice cultivar carrying the corresponding R gene \textit{Pi33}. The \textit{ACE1} encodes a protein with similarity to polyketide synthase (Böhnert \textit{et al.} 2004). Biosynthetic activity of \textit{ACE1} was required for avirulence, supporting the hypothesis that the secondary metabolite synthesized by the \textit{ACE1} is an elicitor that is recognized by the resistant rice cultivar. In the third model, the AVR protein is an enzyme that is translocated into the host cell and interacts with the R protein to trigger HR. One example is the \textit{AVR-Pita} (formerly \textit{Avr2-Yamo}) gene, the first cloned \textit{AVR} gene from \textit{M. grisea}, that governs cultivar specificity by preventing the fungus from infecting the rice cultivar Yashiro-mochi that carries the corresponding \textit{R} gene \textit{Pi-ta} (Valent and Chumley 1994). Because both \textit{Pi-ta} and \textit{AVR-Pita} have been cloned and characterized, studies of a corresponding \textit{AVR-Pita/Pi-ta} gene pair have provided insights into the mechanisms of the biochemical basis of \textit{AVR} gene function (Bryan \textit{et al.} 2000; Jia \textit{et al.} 2000; Orbach \textit{et al.} 2000). \textit{AVR-Pita}, cloned by a map-based cloning strategy from the rice pathogenic strain O-137, encodes a 223 amino acid protein with a signal peptide (Jia \textit{et al.} 2000). The deduced amino acid sequence showed similarity to fungal zinc metalloproteases. \textit{AVR-Pita} may function through its enzymatic activity as a metalloprotease: Site-directed mutations of the putative active sites of the metallopeptidase completely abolished avirulence activity. In addition, it has been shown that the \textit{AVR-Pita} protein (\textit{AVR-Pita}_{223}) is processed to yield a 176 amino acid protein (\textit{AVR-Pita}_{176}), which is the avirulence-active form. This \textit{AVR-Pita}_{176} corresponds to the processed forms of other fungal metalloproteases (Jia \textit{et al.} 2000; Tatsumi \textit{et al.} 1991). The corresponding \textit{R} gene \textit{Pi-ta} gene, cloned by a map-based cloning strategy, encodes a putative cytoplasmic protein of 928 amino acids with a centrally located nucleotide binding site and a carboxyl terminal leucine rich domain (LRD) (Bryan \textit{et al.} 2000). The specific direct interaction between \textit{Pi-ta} and \textit{AVR-Pita} has been demonstrated (Jia \textit{et al.} 2000): In a yeast two-hybrid assay, the putative mature form \textit{AVR-Pita}_{176}, but not the \textit{AVR-Pita}_{223} protein, binds specifically to the LRD of the \textit{Pi-ta} protein. \textit{In vitro} binding studies also supported their specific binding. Interestingly, zinc was required for the interaction, supporting the role of the \textit{AVR-Pita} as a zinc metalloprotease in the interaction with \textit{Pi-ta}. Studies indicate that their interaction occurs inside the host cell:
Application of AVR-Pita\textsubscript{176} into the apoplastic spaces of Pi-ta rice by vacuum infiltration or spray inoculation of AVR-Pita\textsubscript{176} failed to trigger resistance responses including HR. In contrast, using a transient expression system, AVR-Pita\textsubscript{176} was demonstrated to trigger HR when expressed inside the cell of Pi-ta rice, indicating a cytoplasmic site of action. Therefore, it has been proposed that AVR-Pita is translocated from the fungus into the cytoplasm of the host cell where it binds to the LRD region of the Pi-ta to trigger Pi-ta mediated resistance responses (JIA \textit{et al.} 2000).

1.2.4 Role of \textit{AVR} genes for pathogen fitness

Why pathogens carry genes that function to limit their host ranges is an intriguing question. If no evolutionary advantage had been associated with \textit{AVR} genes, such genes would have disappeared from pathogen populations due to selection pressure exerted by resistant hosts. Therefore, intrinsic function(s) that would contribute to pathogen fitness has been proposed to explain the maintenance of \textit{AVR} genes in pathogen populations. Fitness is defined as the ability of an organism to survive and reproduce (CROW 1986). In the case of plant pathogens, fitness can be quantified by measuring components of fitness such as reproductive rate, infection efficiency, and amount of disease (ANTONOVICS and ALEXANDER 1989; LEACH \textit{et al.} 2001). Understanding of the role of \textit{AVR} genes for pathogen fitness is important because the assessment of the fitness penalty associated with the loss of a given \textit{AVR} gene can be an indicator of the durability of the corresponding \textit{R} gene deployed in the crops (BAI \textit{et al.} 2000; LEACH \textit{et al.} 2001). One of the direct ways to test a putative fitness function of an \textit{AVR} gene is to compare the phenotypes \textit{in vitro} and \textit{in planta} between the wild-type carrying the \textit{AVR} gene and a mutant lacking the \textit{AVR} gene. Before cloning technology existed, P.R. Day used three strains of tomato leaf mold \textit{C. fulvum} (e.g., pigmented wild-type strain carrying the \textit{AVR} gene, P-AVR; non-pigmented strain carrying \textit{AVR} gene, NP-AVR; and pigmented strain lacking \textit{AVR} gene, P-avr) and carried out an in planta competition experiment (DAY 1968). He first inoculated tomato seedlings with equal numbers of conidia from the P-AVR strain and the NP-AVR strain. Both of them carried identical \textit{AVR} genes. The
difference in virulence was indirectly determined by harvesting the conidia produced on inoculated seedlings, plating them on medium, and counting the number of colonies of each color. The NP-AVR strain was less virulent than the P-AVR strain. Then, the virulence of the NP-AVR strain was compared to that of the P-avr using the same competition experiment. The NP-AVR strain was more virulent than the P-avr. Results from the competition experiment indicate that the AVR gene contributed to pathogen virulence. However, it cannot be ruled out that the reduced virulence was due to other genetic factors among the strains used. Using cloning technology, more direct and accurate functional analysis was carried out first with the avrBs2 gene from Xanthomonas campestris pv. vesicatoria, the causal agent of bacterial spot disease of pepper and tomato (Kearney and Staskawicz 1990). X. campestris pv. vesicatoria expressing the avrBs2 induces resistance responses including HR on pepper carrying the corresponding R gene Bs2. The strains of X. campestris pv. vesicatoria containing an inactivated avrBs2 exhibited less virulence on susceptible pepper cultivar (lacking Bs2) than the wild-type strain. Sequences homologous to avrBs2 occur in all strains of X. campestris pv. vesicatoria and in many pathovars of X. campestris and X. oryzae, indicating the persistence of the AVR gene in pathogen populations. Mutations in avrBs2 homologs in X. campestris pv. alfaflae also caused reduced virulence on alfalfa. These studies indicate that avrBs2 is required for full virulence. Since then, a role in virulence was subsequently demonstrated for many bacterial AVR genes, including avrA, avrB, avrE, and avrRPM1 from Pseudomonas syringae and pthA, avrB6, and avrXa7 from various Xanthomonas species (Leach and White 1996).

In contrast to bacterial AVR genes, only a few fungal AVR genes have been demonstrated to contribute to pathogen fitness. They include ECP1 and ECP2 (ExtraCellular Protein 1 and 2) from C. fulvum (Laugé et al. 1997) and NIP1 from R. secalis (Rohe et al. 1995). The ECP1 and ECP2, which trigger HR on certain tomato cultivars, were originally characterized as fungal extracellular proteins abundant in the apoplastic space of the host plants infected with C. fulvum. Targeted deletion analyses showed that both of them were required for full virulence, but disruption of ECP2 caused more severe reduction in virulence (Laugé et al. 1997). The NIP1 (also called AvrRrs1)
of *R. secalis* induces resistance responses in barley cultivars containing the corresponding *R* gene *Rrs1* (ROHE *et al.* 1995). Wild-type strains lacking the *NIP1* gene and mutants carrying a mutated *NIP1* exhibited a significant reduction in virulence on barley (ROHE *et al.* 1995). Population genetic analysis showed the coexistence of *R. secalis* strains with and without *NIP1*. This implies that the selective advantage of *NIP1* or the disadvantage by *NIP1* deletion may not be high; otherwise strains with *NIP1* would have become fixed (SCHURCH *et al.* 2004).

Most fungal *AVR* genes are known only as avirulence determinants. For example, *AVR9* disruption did not show any altered phenotype *in vitro* and pathogenicity on tomato compared to isogenic wild-type strain. The failure to detect another function rather than avirulence may be because the experiments were not sensitive enough to recognize small difference in phenotype or because the fitness penalty due to loss of *AVR* gene function might have been compensated by functional redundancy (COLLMER 1998). Small differences in growth rates are difficult to detect in short-term experiments even under *in vitro* conditions. However, a small difference can effectively lead to the selection for *AVR* genes over long periods.

1.2.5 Evolution of *AVR* genes

Pathogens and hosts undergo coevolutionary processes in which reciprocal and adaptive genetic changes occur for plant resistance specificity and pathogen virulence in response to each other. This continuing process is often described as an ‘arms race’ (accumulated improvements in both populations) or ‘Red Queen’ dynamics (running as fast as you can to stay in the same place) (WOOLHOUSE *et al.* 2002). In natural ecosystems, this type of evolution, which guarantees survival of both hosts and their pathogens, exists because of the balance between the emergence of novel resistance specificities and the occurrence of specific virulent pathogens (PRELL and DAY 2001). The genetic polymorphism associated with host resistance provides a potential genetic basis for new resistance specificities. However, in agricultural ecosystems, where genetically uniform cultivars are grown, new pathogen races can arise efficiently in
monoculture, resulting in rapid breakdown of the introduced resistance. Because cultivar resistance is determined by direct or indirect recognition of pathogen $AVR$ gene products by host $R$ gene products, mutations of $AVR$ genes allow pathogens to avoid recognition and overcome resistance. Some $AVR$ genes are known to contribute to pathogen fitness. For such $AVR$ genes, selection would favor the least fitness penalties by any changes. Evidence described below indicates that new virulent pathogens either have inactivated $AVR$ genes (loss of function) or have altered $AVR$ genes so that they lack avirulence activity (divergence of gene function).

1.2.5.1 Loss of $AVR$ gene function

Complete inactivation of $AVR$ genes can occur by several mechanisms, including partial or complete deletion of the gene, frameshift or nonsense mutations, and insertion of transposable elements. Mechanisms that affect the regulation of $AVR$ gene expression may also lead to loss of avirulence function.

In the rice blast fungus $M. grisea$, spontaneous virulent mutants that lack $AVR$-$Pita$ were frequently isolated after inoculating rice carrying the corresponding $R$ gene $Pi-ta$. Some cases of loss in avirulence were due to nonsense mutations of the $AVR$-$Pita$, resulting in truncated proteins, but the majority was due to the deletions ranging from 100 bp to over 10 kb at the $AVR$-$Pita$ locus (ORBACH et al. 2000). $AVR$-$Pita$ is located extremely close to a telomere. Its stop codon is separated from the start of a telomere repeat by only 48 bp. In individual Saccharomyces cerevisiae strains, frequent deletion of some of the subtelomeric loci was observed, indicating that genes located near telomeres are genetically unstable (CARLSON et al. 1985). Similarly, hypervariability of the $AVR$-$Pita$ appears to be associated with its genomic location near a telomere. Another $M. grisea$ $AVR$ gene, $AVR1$-$CO39$, has been found to be deleted partly or entirely in those isolates that are virulent on the rice cultivar CO39 containing $Pi$-$CO39(t)$ (FARMAN et al. 2002). The close association of DNA repeat elements at the point of the $AVR1$-$CO39$ deletion junctions suggests the possible involvement of repeat sequences in the variation the $AVR1$-$CO39$ locus (FARMAN et al. 2002). In the tomato pathogen $C. fulvum$, several
AVR genes have been well characterized, including Avr2, Avr4, and Avr9 that correspond to R genes Cf2, Cf4, and Cf9 (Joosten and de Wit 1999). Virulence on Cf2 tomato resulted from the production of truncated AVR2 proteins that lacked at least three cysteine residues through diverse mutations (Luderer et al. 2002). In the strains that are virulent on Cf4 tomato, various single point mutations were found in the Avr4 gene that result in unstable AVR4 proteins (Joosten et al. 1997). The strains that were virulent on Cf9 tomato lacked the Avr9 gene (van Kan et al. 1991). In the barley pathogen R. secalis, a complete lack of the NIP1 gene was responsible for overcoming Rrs1 resistance (Rohe et al. 1995). Insertions of transposable elements have been found in the open reading frame of the C. fulvum Avr2 gene (Luderer et al. 2002) and M. grisea ACE1 (Böhnert et al. 2004) and the promoter region of the M. grisea AVR-Pita gene (Kang et al. 2001). These insertions were responsible for the loss of avirulence, likely due to interference in gene structure or gene expression.

1.2.5.2 Divergence (Alternation) of AVR gene function

AVR genes would evolve to avoid R gene-mediated recognition while maintaining their intrinsic function, if they contribute to pathogen fitness. Because recognition occurs through specific interactions directly or indirectly between products of AVR and R genes, modifications of protein structure or biochemical activity of AVR genes would elude the R gene products. In pathogen populations, the presence of different alleles of an AVR gene that encode functional proteins but lack avirulence activity would be indirect evidence for such evolution. For example, some field isolates of M. grisea have a virulent allele of the AVR-Pita gene that failed to confer avirulence. The encoded protein has a single amino acid substitution in the protease motif relative to the AVR-Pita protein that confers avirulence. Because the changed amino acid still fits the defined protease motif, the virulent AVR-Pita is likely to function as a metalloprotease (Montenegro-Chamorro 1997; Orbach et al. 2000). It remains to be demonstrated if the virulent AVR-Pita indeed has the protease activity and contributes to pathogen fitness, if it does at all.
Novel gene function can arise through gene duplication and subsequent functional divergence of duplicated genes (Ohno 1970). Immediately after duplication, two genes have identical sequences and functions. As different sequence changes accumulate in each of these genes, their functions will also diverge with time. One copy of the gene can retain the original function while the duplicate becomes nonfunctional or gains a new function. Recent studies have shown that gene duplication may also provide a means for preserving function in that even after two copies of a gene have diverged widely, they can still substitute for each other’s functions at some degree (Gu et al. 2003). All the amplified copies of a gene that can be recognized by their significant similarity (at least 50% similarity at the amino acid level) are considered members of the same gene family (Dayhoff 1978). *M. grisea* has the PWL gene family that governs host specificity (Kang et al. 1995; Sweigard et al. 1995). Two family members PWL1 and PWL2 (75% amino acid similarity each other) were functional in preventing the infection of weeping lovegrass. The other members PWL3 and PWL4 (51% and 57% amino acid similarity to the PWL2, respectively) were nonfunctional. Interestingly, when the PWL3 open reading frame (ORF) or PWL4 ORF was placed under the control of either PWL1 promoter or PWL2 promoter, PWL4 became functional, but PWL3 remained nonfunctional, suggesting that PWL4 is nonfunctional because of improper expression, while PWL3 is nonfunctional because it does not encode a protein that confers avirulence (Kang et al. 1995). A possible interpretation for the nonfunctional PWL3 and PWL4 is that these genes would have lost avirulence activity as a result of the functional divergence through accumulation of mutations in ORF and promoter, respectively, after gene duplication. Although any intrinsic functions other than avirulence are not yet known, the duplications and persistence of the family members in pathogen populations imply the existence of a putative beneficial function of the PWL genes.

Positive selection for diversification may act on the evolution of the R genes and AVR genes, because these genes are expected to continually adapt in response to each other. Diversifying (positive) selection can be inferred through molecular evolutionary analysis. Nonsynonymous amino acid changes (N) are more likely to change the function of a protein than synonymous amino acid changes (S). If diversifying selection acts on
the evolution of a protein, the rate of N should exceed that of S. This is because advantageous mutations undergo fixation more rapidly than neutral mutations. Therefore, one way to detect the existence of diversifying selection is to show that the number of nonsynonymous substitutions per nonsynonymous site (dN) is significantly greater than the number of synonymous substitutions per synonymous site (dS) (Nei and Kumar 2000). In host R gene products, leucine-rich repeat (LRR) region has been demonstrated as the major determinants of recognition specificity for AVR gene products (Dangl and Jones 2001). Evolutionary analysis has shown that dN is significantly higher than dS in the LRR region of several R genes, reflecting diversifying selection (Bergelson et al. 2001; Mondragón-Palomino et al. 2002). This suggests that R genes evolve to improve the recognition specificities against AVR genes. In the barley pathogen R. secalis, the NIP1 gene that encodes the protein having both avirulence activity and toxicity has been shown to be under diversifying selection (Schurch et al. 2004). Some mutations in NIP1 resulted in abolishing avirulence activity on barley that carries the corresponding R gene Rrs1 and may have been retained by diversifying selection acting on the NIP1. Strong diversifying selection has been also shown for the AVR gene AvrL567 of flax rust fungus M. lini (Dodds et al. 2004). Comparison of three AvrL567 genes (AvrL567-A, -B, and –C) showed an excess of nucleotide variation within the coding sequence relative to flanking sequences, suggesting that selection has contributed to the fast evolution of the coding sequences. In addition, dN is significantly higher than dS. The Avr567-C gene, which failed to confer avirulence, had amino acid differences from Avr567-A and –B, possibly driven by diversifying selection. The L gene of Linum usitatissimum, which is the corresponding R gene of the AvrL567, is also shown to have been under diversifying selection (Dodds et al. 2000). The observation that both AvrL567 and L genes are driven by diversifying selection provides an example of a coevolutionary arms race between R and AVR genes for host resistance and pathogen virulence.
1.3 Conclusion

*R* gene-mediated resistance has been widely utilized for controlling plant diseases. However, one of the challenges in this approach is the durability of resistance, which is often defeated by changes in pathogen *AVR* genes. Therefore, better understanding of the function and variation of pathogen *AVR* genes will contribute to developing durable plant resistance. In this introductory chapter, the current status of knowledge on the function and variation of pathogen *AVR* genes was reviewed. An intriguing question of why pathogens carry *AVR* genes that potentially limit their host range has been addressed by studies that demonstrate the dual functions of *AVR* genes in avirulence as well as pathogenicity. Diverse mutations in *AVR* genes allow pathogens to avoid *R* gene-mediated recognition, therefore overcoming resistance. Thus, durable resistance may be achieved by utilizing plant *R* genes that can recognize pathogen *AVR* genes that are essential for the pathogen itself.

1.4 Outline of the thesis

The long-term goal of this study is to understand the mechanisms by which fungal *AVR* genes have evolved. Because pathogens evade *R* gene-mediated resistance by changing corresponding *AVR* genes, as shown in a number of pathogen systems, a better understanding of the mechanisms underpinning pathogen *AVR* gene variation will contribute to developing a novel strategy for durable resistance.

Toward such an understanding, this study focuses on characterizing the *AVR-Pita* gene family, an *AVR* gene family in *Magnaporthe grisea* that frequently mutates, as a model. Specific objectives of this thesis research are as follows:

1. determine the distribution patterns of *AVR-Pita* and its homologs among isolates of *M. grisea* from diverse hosts
2. characterize the function of the *AVR-Pita* homologs
3. determine the genetic mechanisms responsible for the evolution of the *AVR-Pita* gene family
In Chapter 2, the data presented indicate that AVR-Pita is a member of a new gene family (termed MgMEP) consisting of at least three family members. The three members (termed MgMEP1, MgMEP2, and MgMEP3) were characterized for their distribution patterns in M. grisea populations, function, and evolutionary history. In Chapter 3, data on the molecular mechanisms responsible for the evolution of the MgMEP gene family are described. Chapter 4 describes a novel tool, based on Agrobacterium tumefaciens mediated transformation, for manipulating fungal genes. A detailed protocol for this tool is described in the Appendix. Finally, in Chapter 5, overall conclusions of this study, discussions on the implications of this study and future directions are presented.
Chapter 2

CHARACTERIZATION OF THE MGMEP (FORMERLY AVR-PITA) AVIRULENCE GENE FAMILY AMONG ISOLATES OF MAGNAPORTHE GRISEA FROM DIVERSE HOSTS

2.1 Introduction

Rice blast disease, one of the most devastating crop diseases worldwide, is caused by the haploid heterothallic ascomycete Magnaporthe grisea (Herbert) Barr (OU 1985). This fungus has a wide host range that includes more than 50 grass species (OU 1985). However, individual isolates are limited to parasitize only a few host species. Among isolates that infect rice (Oryzae sativa), hundreds of pathogen races have been identified based on their virulence or avirulence on a set of rice cultivars (Ziegler et al. 1994). Control of rice blast disease by breeding resistant cultivars has not been durable, typically due to frequent appearance of new races that are virulent to previously resistant cultivars.

The rice blast pathosystem is a classical gene-for-gene host-pathogen system, in which M. grisea avirulence (AVR) genes trigger resistant responses in host expressing corresponding resistance (R) genes (Orbach et al. 2000; Ziegler et al. 1994). In M. grisea, several AVR genes (AVR-Pita, AVR1-CO39, PWL1, PWL2, and ACE1) have been cloned (Böhnert et al. 2004; Farman et al. 2002; Kang et al. 1995; Orbach et al. 2000; Sweigard et al. 1995) and these AVR genes prevent the fungus from infecting certain host species or rice cultivars in a gene-for-gene manner. It has been an intriguing question why a pathogen carries an AVR gene that appears only to limit its virulence. Recent studies, however, have shown that AVR genes can function as virulence factors against certain host genotypes that do not carry corresponding R genes (Leach et al. 2001). A number of bacterial AVR genes have been shown to be necessary for the fitness of the pathogen (Leach and White 1996). Some of the fungal AVR genes have also been demonstrated to contribute to pathogen fitness. They include ECP1 and ECP2
(ExtraCellular Protein 1 and 2) from *Cladosporium funvum* (LAUGÉ et al. 1997) and NIP1 from *Rhynchosporium secalis* (ROHE et al. 1995).

*AVR-Pita*, one of the *AVR* genes in *M. grisea*, was cloned from the rice pathogenic isolate O-137 (ORBACH et al. 2000). This gene encodes a putative 223 amino acid protein with similarity to fungal zinc metallopeptidases. Its metallopeptidase activity appeared critical in conferring avirulence activity. The AVR-Pita protein contains a signal sequence for extracellular secretion and appears to be processed to a mature protein of 176 amino acids. The mature AVR-Pita protein was demonstrated to interact directly with the corresponding *R* gene product, Pi-ta, which is a predicted cytoplasmic receptor protein with a centrally located nucleotide binding site and a carboxyl terminal leucine rich domain.

Occurrence of the *AVR-Pita* gene in field isolates from rice has been determined by DNA gel blot analysis using the O-137 *AVR-Pita* gene as a probe. Restriction fragment length polymorphisms (RFLPs) associated with *AVR-Pita* gene were conserved within single genetic lineages in Colombia and in the Philippines but generally varied between lineages (MONTENEGRO-CHAMORRO 1997; ZEIGLER et al. 1995). Certain RFLPs correlated with the avirulence phenotype, while others were associated with the virulence phenotype. Sequence analysis of *AVR-Pita* alleles showed limited variation of amino acid sequences within alleles from avirulent or virulent lineages. Considering the frequent spontaneous mutations of the O-137 *AVR-Pita* gene into a virulent allele, resulting in gain of virulence on Pi-ta rice, the strong conservation of *AVR-Pita* in field isolates is noteworthy. The high level of variation in the O-137 *AVR-Pita* gene is likely due to the telomeric location of the gene, of which the stop codon is separated from telomeric repeat sequences only by 48 bp (ORBACH et al. 2000). Genetic instability is associated with telomeric location in other organisms (CARLSON et al. 1985; CHARRON and MICHELS 1988; HEATHER and TRASK 2002). Thus, the conserved *AVR-Pita* alleles in field isolates may be due to their stable chromosomal locations, unlike to the telomeric O-137 *AVR-Pita*. Alternatively, regardless chromosomal locations, *AVR-Pita* may have been maintained possibly because of its virulence function.
In this work, the occurrence of the *AVR-Pita* gene was further surveyed among isolates of *M. grisea* from diverse hosts. This survey revealed that *AVR-Pita* is a member of a new gene family, *MgMEP*, which is present not only in rice-infecting isolates, but also in other grass-infecting isolates. Members of this gene family were cloned and characterized with respect to functional divergence, phylogenetic relationships with O-137 *AVR-Pita*, and their occurrence in *M. grisea* populations.

2.2 Materials and Methods

2.2.1 Strains and growth conditions

The *M. grisea* field isolates and laboratory strains used in this study are described in Table 1. *Escherichia coli* strain XL1-blue MRF’ was used for maintaining plasmids. *Agrobacterium tumefaciens* strain EHA105 (KLEE 2000) was used to transform *M. grisea* strain CP987 (SWEIGARD et al. 1995). All *M. grisea* isolates were grown and maintained on oatmeal agar medium as previously described (VALENT et al. 1991).
Table 1. *Magnaporthe grisea* field isolates used in this studya

<table>
<thead>
<tr>
<th>Strain</th>
<th>Host b</th>
<th>Country of origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>O-135</td>
<td><em>Oryza sativa</em></td>
<td>China</td>
</tr>
<tr>
<td>O-137</td>
<td><em>Oryza sativa</em></td>
<td>China</td>
</tr>
<tr>
<td>Guy11</td>
<td><em>Oryza sativa</em></td>
<td>Guyana</td>
</tr>
<tr>
<td>G-160</td>
<td><em>Leersia hexandra</em></td>
<td>Philippines</td>
</tr>
<tr>
<td>O-315</td>
<td><em>Oryza longistaminata</em></td>
<td>Cameroon</td>
</tr>
<tr>
<td>T-5</td>
<td><em>Triticum cv. Batuira</em></td>
<td>Brazil</td>
</tr>
<tr>
<td>G-158</td>
<td><em>Triticum X Secale</em></td>
<td>Brazil</td>
</tr>
<tr>
<td>G-200</td>
<td><em>Eleusine indica</em></td>
<td>Ivory Coast</td>
</tr>
<tr>
<td>G-77</td>
<td><em>Eleusine indica</em></td>
<td>Philippines</td>
</tr>
<tr>
<td>G-17</td>
<td><em>Eragrostis curvula</em></td>
<td>Japan</td>
</tr>
<tr>
<td>G-22</td>
<td><em>Eleusine coracana</em></td>
<td>Japan</td>
</tr>
<tr>
<td>G-176</td>
<td><em>Eleusine coracana</em></td>
<td>Nepal</td>
</tr>
<tr>
<td>G-219</td>
<td><em>Panicum maximum</em></td>
<td>India</td>
</tr>
<tr>
<td>G-165</td>
<td><em>Brachiaria distachya</em></td>
<td>Philippines</td>
</tr>
<tr>
<td>G-218</td>
<td><em>Panicum coloratum</em></td>
<td>Japan</td>
</tr>
<tr>
<td>G-220</td>
<td><em>Panicum repens</em></td>
<td>India</td>
</tr>
<tr>
<td>G-221</td>
<td><em>Pennisetum clandestinum</em></td>
<td>Japan</td>
</tr>
<tr>
<td>G-156</td>
<td><em>Digitaria horizontalis</em></td>
<td>Brazil</td>
</tr>
<tr>
<td>G-11</td>
<td><em>Digitaria sanguinalis</em></td>
<td>Japan</td>
</tr>
<tr>
<td>G-1</td>
<td><em>Digitaria sanguinalis</em></td>
<td>USA</td>
</tr>
<tr>
<td>G-213</td>
<td><em>Digitaria smutii</em></td>
<td>Japan</td>
</tr>
<tr>
<td>G-32</td>
<td><em>Digitaria spp.</em></td>
<td>Japan</td>
</tr>
<tr>
<td>G-189</td>
<td><em>Digitaria spp.</em></td>
<td>USA</td>
</tr>
<tr>
<td>G-123</td>
<td><em>Pennisetum glaucum</em></td>
<td>USA</td>
</tr>
<tr>
<td>G-78</td>
<td><em>Pennisetum polystachyon</em></td>
<td>Philippines</td>
</tr>
<tr>
<td>G-223</td>
<td><em>Pennisetum typhoideum</em></td>
<td>Burkina Faso</td>
</tr>
<tr>
<td>G-194</td>
<td><em>Leersia oryzoides</em></td>
<td>Japan</td>
</tr>
<tr>
<td>G-229</td>
<td><em>Cyperus brevifolius</em></td>
<td>Philippines</td>
</tr>
<tr>
<td>G-231</td>
<td><em>Cyperus rotundus</em></td>
<td>Philippines</td>
</tr>
</tbody>
</table>

*aAdditional information about these isolates can be found in Kang et al. (1995) and Sweigard et al. (1995)*

b*The host plant on which the isolate was found.*
2.2.2 Cloning of the MgMEP homologs

2.2.2.1 Cloning by genomic DNA library screening

Clones containing the AVR-Pita homologs (named MgMEP1, MgMEP2, and MgMEP3) were isolated from genomic DNA libraries of M. grisea field isolates G-1, G-17, G-22, G-78, G-213, and G-223, constructed using the λGEM12 vector (Promega, Madison, WI). Genomic DNA libraries were screened by hybridization under low-stringency conditions with a $^{32}$P-labeled AVR-Pita probe that was prepared as follows: A 0.9 kb PCR product, amplified from M. grisea O-137 using the primers LF9 (5’-GCGATTTCGGCCTTCACC-3’) and Pita1 (5’-CCCTCTATTGTTAGATTGA-3’), was cloned into pGEM-T Easy vector (Promega, Madison, WI). The resulting clone was digested with SspI to liberate 0.5 kb fragment to be used for random priming with Klenow fragment and $\alpha ^{32}$P-labeled dCTP. The resulting positive plaques were purified and the λ DNAs were isolated using the plate lysate method (SAMBROOK et al. 1989). Electrophoresis of enzyme-digested lambda DNA and Southern blotting were performed according to standard protocols (SAMBROOK et al. 1989). Restriction fragments hybridizing to the probe were subcloned into pGEM3Zf (Promega, Madison, WI) and sequenced utilizing a combination of the EZ::TN transposon insertion system (Epicentre Technologies, Madison, WI), further subcloning, and primer walking.

2.2.2.2 Cloning by PCR

17 MgMEP genes were amplified from 14 M. grisea field isolates using PCR. The primers (Table 2) were designed based on the conserved regions of the MgMEP1 genes from O-137, G-1, G-78, and G-223, the MgMEP2 gene in G-213, and the MgMEP3 gene
from G-17 and G-22. PCR amplifications were performed using the FailSafe PCR system (Epicentre Technologies, Madison, WI). The PCR cycling program included an initial denaturation for 5 min at 95°C, 30 cycles of 30 sec denaturation at 95°C, 30 sec annealing at 58°C, 1-min extension at 72°C, and followed by a final extension for 10 min at 72°C. PCR products were cloned in a pGEM-T Easy vector (Promega, Madison, WI). Three individual PCR clones were included in DNA sequence analysis to identify sequence changes introduced during PCR amplification.

**Table 2. Primers used in this study**

<table>
<thead>
<tr>
<th>Product</th>
<th>Primer</th>
<th>Sequence (5’→ 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MgMEP1</td>
<td>Fun1</td>
<td>ATAAAATCGAC (C/G) CGTTTCCG</td>
</tr>
<tr>
<td></td>
<td>Fun2-1</td>
<td>GATTCCCTCCATTCAACAC</td>
</tr>
<tr>
<td>MgMEP2</td>
<td>Fun1</td>
<td>Same as the Fun1 above</td>
</tr>
<tr>
<td></td>
<td>Fun2-2</td>
<td>GCCGTGACAAATCCCTTTAT</td>
</tr>
<tr>
<td>MgMEP3</td>
<td>nFun1</td>
<td>TTTCCCTCTCTTTTCGTTTTTTC</td>
</tr>
<tr>
<td></td>
<td>nFun2</td>
<td>TGTCGTTCTCTCTGGCTTCC</td>
</tr>
<tr>
<td>P&lt;sub&gt;MgMEP1&lt;/sub&gt;</td>
<td>Fs-P1</td>
<td>CGGAATTTCGCCGAGTCGTTCTGA</td>
</tr>
<tr>
<td></td>
<td>Fs-P2</td>
<td>CGGGATCTGTTAATTGTGCGAAGTTTTTT</td>
</tr>
<tr>
<td>C&lt;sub&gt;MgMEP1&lt;/sub&gt;</td>
<td>Fs-C1</td>
<td>CGGGATCCTTGCAATTATGCTTTTTTTATTC</td>
</tr>
<tr>
<td></td>
<td>Fs-C2</td>
<td>CATGCATGCCCCTCTCTGTGGATGTTG</td>
</tr>
<tr>
<td>P&lt;sub&gt;MgMEP2&lt;/sub&gt;</td>
<td>Fw-P1</td>
<td>CGGAATTCTACGGCATATTGTTTGGTTATTC</td>
</tr>
<tr>
<td></td>
<td>Fw-P2</td>
<td>CGGGATCTGCTATGGCGAATTGTGCAAAAGTTTTTT</td>
</tr>
<tr>
<td>C&lt;sub&gt;MgMEP2&lt;/sub&gt;</td>
<td>Fs-C1</td>
<td>Same as the Fs-C1 above</td>
</tr>
<tr>
<td></td>
<td>Fw-C2</td>
<td>CATGCATGCCTGGCCGGATGACATTTC</td>
</tr>
<tr>
<td>P&lt;sub&gt;MgMEP3&lt;/sub&gt;</td>
<td>nF-P1</td>
<td>CGGAATTCAAGTAAGCAGAAAATACAG</td>
</tr>
<tr>
<td></td>
<td>nF-P2</td>
<td>CGGGATCCTACAATTGTGCGAGATTTTTT</td>
</tr>
<tr>
<td>C&lt;sub&gt;MgMEP3&lt;/sub&gt;</td>
<td>nF-C1</td>
<td>CGGGATCCCGCAATTATGCTTTTTTTATAT</td>
</tr>
<tr>
<td></td>
<td>nF-C2</td>
<td>CATGCATGCCCCTCCAAACAAAAGGG</td>
</tr>
<tr>
<td>ITS</td>
<td>ITS1</td>
<td>TCCGTTAGGTGAACCTGCGG</td>
</tr>
<tr>
<td></td>
<td>ITS4</td>
<td>TCCTCCGCTTATTGATATGC</td>
</tr>
</tbody>
</table>
2.2.2.3. Cloning *in silico*

The *M. grisea* strain 70-15 is a laboratory strain that has been chosen for genome sequencing (http://www.broad.mit.edu/annotation/fungi/magnaporthe/index.html). This strain was derived from a cross between the rice infecting strain Guy 11 and a weeping lovegrass infecting strain followed by numerous backcrosses to Guy 11 (CHAO and ELLINGBOE 1991; LEUNG et al. 1988). Genomic sequences containing the *MgMEP1* (locus ID MG11081) and *MgMEP3* (no locus ID assigned) genes were identified by a BLASTN search of the genome database with the O-137 *AVR-Pita* gene as a query sequence.

2.2.3 DNA sequencing and analysis

DNA sequencing reactions were performed using the ABI Prism Big-Dye terminator sequencing kit (Applied Biosystems, Foster City, CA) and analyzed with an ABI377 sequencer. DNA reads were edited manually for ambiguous sites and assembled using the program Autoassembler (Applied Biosystems, Foster City, CA). Signal peptides were predicted using SignalP v.3.0 (http://www.cbs.dtu.dk/services/SignalP/). Peptidase units of MgMEP were identified using the resource MEROPS (http://merops.sanger.ac.uk/) and comparing MgMEP members with the type members of the M35 family.

2.2.4 Southern analysis

Genomic DNA from fungal strains was digested with *EcoRI*, subjected to electrophoresis on a 0.8% agarose gel and transferred to Hybond N+ membrane (Amersham, Piscataway, NJ) according to the manufacturer’s instructions. The membranes were hybridized at 65°C overnight with ³²P-labeled probes prepared by random priming. After hybridization, the membranes were washed twice in 2×SSPE and
0.1% SDS for 15 min at 65°C followed by two additional washes in 0.1× SSPE and 0.1% SDS for 15 min at 65°C and exposed to Kodak X-ray film.

2.2.5. Sequence alignment and phylogenetic analysis

All sequence data were aligned using the CLUSTALW program (Thompson et al. 1994) with default parameters. Neighbor-joining trees were constructed using the MEGA2 program (Ver. 2.1) (Kumar et al. 2001) with the following parameters: complete deletion of gaps, Kimura 2-parameter model, both transitions and transversions substitutions included, and 1000 bootstrap replicates.

The internal transcribed spacer (ITS) region (including part of the 18S rRNA and 26S rRNA genes and the complete sequence of 5.8S rRNA gene and ITS1 and ITS2) was amplified from 23 M. grisea isolates (Fig. 1) with the primers ITS1 and ITS4 (Table 2) as previously described (White et al. 1990). PCR amplifications were performed using the FailSafe PCR system (Epicentre Technologies, Madison, WI). PCR cycling program included an initial denaturation for 2 min at 94°C, 30 cycles of 1 min denaturation at 94°C, 1 min annealing at 55°C, 1 min extension at 72°C, followed by a final extension for 10 min at 72°C. PCR products were purified with QIAGEN spin columns and directly sequenced on both strands with the same primers used in the PCR amplification. The gene tree of 23 MgMEP genes was constructed as described above.

2.2.6 Construction of chimeric MgMEP genes

The primers used to amplify the promoters (P_MgMEP1, P_MgMEP2, and P_MgMEP3) and genomic coding sequences (gCDS; C_MgMEP1, C_MgMEP2, and C_MgMEP3) of the MgMEP1, MeMEP2, and MgMEP3 genes are shown in Table 2. These primers contain a restriction enzyme site at the 5' end of each primer to facilitate subsequent cloning: EcoRI and BamHI sites at the 5' end and 3' end for the promoter and BamHI and SphI sites at the 5'
end and 3’ end for gCDS. The location of putative promoter region of each member was defined relative to the translation start site (A of ATG as +1). The promoter fragments for the *MgMEP1*, *MeMEP2*, and *MgMEP3* genes include the sequences of –9 to –489 (481bp), –9 to –650 (642bp), and –9 to –573 (565bp), respectively. The last six nucleotides at the 3’ end of the promoter were altered from 5’-TTATTT-3’ to 5’-GGATCC-3’ to introduce a *Bam*HI site. PCR amplifications were carried out using the FailSafe PCR system (Epicentre Technologies, Madison, WI). The PCR cycling program included an initial denaturation for 5 min at 95°C, two cycles of 30 sec denaturation at 95°C, 30 sec annealing at 44°C, 1 min extension at 72°C and 25 cycles of 30 sec denaturation at 95°C, 30 sec annealing at 56°C, 1 min extension at 72°C, followed by a final extension for 10 min at 72°C. The gCDS of each member was amplified under the same conditions described above. PCR products were isolated from a gel using QIAGEN spin columns and were cloned in pGEM-T Easy vector (Promega, Madison, WI). All clones were verified by DNA sequencing. Six combinatorial chimeric constructs were generated by cloning an *Eco*RI-*Bam*HI fragment of a promoter (P) and a *Bam*HI-*Sp*HI fragment of a gCDS (C) into the *Eco*RI-*Sp*HI sites of pSK1213, which contains *N. crassa* terminator sequences (T; 300pb) in the *Sp*HI-*Hin*dIII sites of pGEM-3Zf (Promega, Madison, WI). From the resulting clones, a *Eco*RI-*Hin*dIII fragment consisting of P::C::T was isolated and cloned between the *Eco*RI and *Hin*dIII sites of pBHt2 (MULLINS et al. 2001). For those that contain either *MgMEP1* or *MgMEP2* promoter, a partial *Hin*dIII digestion following a complete *Eco*RI digestion was carried out to clone them in pBHt2 due to the presence of a *Hin*dIII site in their promoter sequences.

2.2.7 Fungal transformation

Fungal transformants were generated using the *Agrobacterium tumefaciens*-mediated transformation method described in Appendix. To test whether these transformants harbor an intact chimeric gene, genomic DNAs were extracted from liquid cultures of
individual transformants and subjected to PCR analysis with the primers that amplify the sequences from promoter to terminator in each chimera.

2.3.8 Rice cultivar, fungal transformants, and infection assay

Rice cultivars YT14 (Pi-ta/Pi-ta) and YT16 (pi-ta/pi-ta) were obtained from Dr. Barbara Valent at Kansas State University. YT14 and YT16 were derived from parental cultivars Yashiro-mochi (Pi-ta/Pi-ta), the source of the cloned Pi-ta gene, and Tsuyuake (pi-ta/pi-ta), containing a japonica susceptibility pi-ta allele (Bryan et al. 2000; Jia et al. 2003).

Infection assays were performed as described in (Valent et al. 1991). Rice seeds were treated with 1% NaClO for 20 min for surface sterilization, followed by three washes with sterile water. The treated seeds were germinated in sterile water for two days at room temperature. Five seedlings were transplanted to plastic pots (5-inch in diameter) filled with Metro-mix 360 (The Scotts Company, Marysville, Ohio). Plants were grown in a greenhouse at 25°C to 30°C with a daily cycle of 16 hr of light and 8 hr of darkness. Plants were allowed to grow about two weeks until they reach the four-leaf stage. Five or six independently isolated fungal transformants for each chimeric gene were included in the infection assay. The transformants were grown at 25°C on oatmeal agar plates to produce conidia. Conidia were collected from 10 day old cultures by scraping with 0.01% Tween 20, followed by filtration of conidial suspension through two layers of cheesecloth to remove mycelia and other debris. The concentration of conidia was adjusted to 5 x 10^4 conidia/ml. Five ml of spore suspension were applied per pot using an artist’s airbrush. Plants were kept inside a sealed plastic bag to maintain high humidity for 24 hrs. The plants were then removed from the bags and maintained on a greenhouse bench holding two to three inches of water. For the control experiment, plants were sprayed with 0.01% Tween 20 or inoculated with the avirulent Chinese field isolate O-137 (carrying MgMEPl) and a spontaneous mutant of O-137, CP987 (mgmep1), that became virulent to YT14 (Bryan et al. 2000).
2.3 Results

2.3.1 *AVR-Pita* is a member of a gene family

More than 300 *M. grisea* strains isolated from diverse hosts and different geographical areas were surveyed for the presence of genes homologous to *AVR-Pita* (see Fig. 1A for selected examples). A blot containing *EcoRI*-digested genomic DNA of 29 isolates was hybridized under stringent conditions using the *AVR-Pita* gene as a probe (Fig. 1A, top panel). Fungal strains included in this figure comprise three distinct genetic lineages, represented by those isolated from rice (*Oryzae sativa*), *Digitaria* spp., or *Cyperus* spp. (Fig. 1B). The presence of a hybridizing band was not limited to a particular lineage and the number of hybridizing bands varied among strains, ranging zero to three. No hybridization bands in certain strains is most likely due to a deletion event of the corresponding sequence, because other strains in distantly related lineages had a hybridization signal. A gene duplication event was evidenced by the presence of three bands of equal hybridization intensity in O-315. Multiple bands in Southern blot analysis do not always indicate the presence of multiple copies of a gene, because even a single copy gene can produce more than one hybridizing band if there is an *EcoRI* site within the gene. One of the copies in O-315 was sequenced and found to contain no *EcoRI* site, confirming the presence of multiple copies of the gene in O-315. Some strains had a faint band (G-220, G-213, and G-32 in Fig. 1A, top panel), indicating the presence of sequences that are distantly related to the *AVR-Pita* gene. Two strains isolated from *Cyperus* spp. (G-229 and G-231 in Fig. 1A, top panel), which are the most distantly related to the other *M. grisea* strains (Fig. 1B), showed no hybridization signal.
Figure 1. Distribution pattern of the MgMEP gene family in Magnaporthe grisea isolates from diverse hosts. A, Genomic DNA was digested with EcoRI, electrophoresed, blotted, and probed sequentially with O-137 MgMEP1 (formerly AVR-Pita), G-213 MgMEP2, and G-22 MgMEP3, from the top to bottom, respectively. The previous probe was stripped from the blot before the blot was reprobed. B, Phylogenetic relationships among 23 isolates were inferred from the internal transcribed spacer (ITS) region using the program MEGA2. The Neighbor-joining tree was constructed under the Kimura 2-parameter model with complete deletion of gaps and both transitions and transversions substitutions included. The numbers near the nodes indicate percentage of 1000 bootstrap replicates in which the node was present. Presence or absence of hybridization band for the MgMEP1, MgMEP2, and MgMEP3 (rectangle, triangle, and circle, respectively) in a M. grisea isolate is indicated with a solid-lined or broken-lined shape, respectively. A filled shape indicates that the corresponding gene was cloned and sequenced in this study. A circle filled with lines indicates a pseudogene resulting from frameshift or nonsense mutations.
2.3.2 *AVR-Pita* and its homologs belong to a new peptidase gene family

To clone *AVR-Pita* homologs, genomic DNA libraries of six *M. grisea* field isolates were screened by hybridization using the *AVR-Pita* gene as a probe. These strains were isolated from *Digitaria* spp. (G-1 and G-213), *Pennisetum* spp. (G-78 and G-223), *Eragrostis* (G-17), or *Eleusine* (G-22) (Table 1). Low-stringency hybridization conditions were used to allow the probe to detect even divergent sequences. Positive λ DNAs were purified and subjected to restriction enzyme analysis. DNA fragments that hybridized to the probe were subcloned into pGEM3Zf. DNA sequence analysis of these clones identified the genes that exhibited different degrees of sequence identity (71% - 98% identity; Table 3) to the O-137 *AVR-Pita* gene. The *M. grisea* strain 70-15 is a laboratory strain whose genome has been sequenced. BLAST search of its genome sequence with the O-137 *AVR-Pita* as a query identified two genes of high similarity. One (locus ID MG11081) exhibited 99% DNA sequence identity to the O-137 *AVR-Pita* gene, while the other one (no locus ID assigned) exhibited 70% DNA sequence identity to the query, but 97% DNA sequence identity to those from G-17 and G-22.

| Table 3. Pairwise sequence comparison between *MgMEP* genes<sup>a</sup> |
|-----------------------------|------------------|------------------|------------------|-------------|-------------|-------------|
| Isolates | MgMEP1 O-137 | MgMEP1 G-78 | MgMEP1 G-223 | MgMEP1 G-1 | MgMEP1 G-213 | MgMEP1 G-22 |
| *MgMEP1* | O-137 | 93 | 95 | 94 | 82 | 56 |
| G-78 | 98 | 97 | 98 | 84 | 58 |
| G-223 | 98 | 99 | 99 | 83 | 58 |
| G-1 | 98 | 99 | 99 | 83 | 58 |
| *MgMEP2* | G-213 | 92 | 92 | 91 | 91 | 54 |
| *MgMEP3* | G-22 | 71 | 71 | 71 | 71 | 70 |
| G-17<sup>b</sup> | 71 | 71 | 71 | 71 | 70 | 97 |

<sup>a</sup>The percent identity of DNA sequence (bottom left-hand corner) and percent similarity of amino acid sequence (top right-hand corner).

<sup>b</sup>Protein sequence is not available because of a premature stop codon.
The O-137 \textit{AVR-Pita} gene consists of four exons separated by three introns, encoding a 223 amino acid protein (Orbach \textit{et al.} 2000). The eight genes from seven strains described above (six from field isolates and two from 70-15) had good matches to the exons and introns of the O-137 \textit{AVR-Pita} gene. Conceptual translations were made to obtain amino acid sequences encoded by these eight genes. Most genes were successfully translated to a protein of 224 or 226 amino acids, but one from the strain G-17 had a frameshift mutation near the start codon leading to a nearby stop codon, resulting in a truncated protein (18 amino acids). The amino acid sequences deduced from these eight genes, except one from G-17, showed a high percentage of sequence similarity to O-137 AVR-Pita, ranging from 55% to 82% (Table 3). The O-137 AVR-Pita protein and all its homologs belong to the metallopeptidase family M35 (deuterolysin family, clan MA(M)), which is evidenced by the presence of conserved peptidase motif HEXXH (http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF02102). These gene products are most similar to \textit{Aspergillus oryzae} neutral protease II (NpII), which is a type I metallopeptidase of the M35 family (http://merops.sanger.ac.uk/). Recently, the O-137 AVR-Pita protein was demonstrated to have peptidase activity (Barbara Valent at Kansas State, Personal communication). On the basis of significant sequence similarity with the known metallopeptidases as well as enzymatic activity, the O-137 \textit{AVR-Pita} gene and its homologs from G-1, G-78, G-223, and 70-15 that are 99% identical to \textit{AVR-Pita} were designated as \textit{MgMEP1} (\textit{Magnaporthe grisea Metallopeptidase 1}). The \textit{AVR-Pita} homolog from G-213 was designated as \textit{MgMEP2}, while ones from G-17, G-22 and 70-15 were called \textit{MgMEP3}. Together, these genes form a new gene family (designated as \textit{MgMEP} family) in \textit{M. grisea}.

The EcoRI-digested genomic DNA blot used for the survey of the \textit{MgMEP1} (\textit{AVR-Pita}) homologues was stripped and reprobed with \textit{MgMEP2} and \textit{MgMEP3} (Fig. 1A). Because of the significant sequence identity between \textit{MgMEP1} and \textit{MgMEP2} (92% identity), they cross-hybridized. However, \textit{MgMEP2} picked up sequences that had hybridized weakly to \textit{MgMEP1} or failed to hybridize in some strains (G-220, G-213, G-32, and G-189). The presence of additional genes that hybridize to \textit{MgMEP2}, but not to \textit{MgMEP1}, suggests the presence of additional members in the \textit{MgMEP} gene family. The
MgMEP3 hybridization profiles substantially differed from those by MgMEP1 or MgMEP2. Its existence appeared to be limited to a certain lineage of rice-infecting isolates and other closely related isolates (first 17 lanes from the left in Fig.1A, bottom). The presence of a single hybridization band indicates that there has not been gene duplication of the MgMEP3 gene, which differs from the MgMEP1 and MgMEP2 genes that appeared to have been duplicated in certain strains.

2.3.3 Functional characterization of members of the MgMEP family

The MgMEP1 genes from G-1, G-78, G-223 and MgMEP2 from G-213 conferred avirulence when transformed into strain CP987 (mgmep1) that is virulent to Pi-ta rice cultivar YT-14 (Pi-ta/Pi-ta). In contrast, the MgMEP1 gene cloned from TH3 (a rice pathogen from Thailand) and Guy11 (a rice pathogen from French Guiana) lacked avirulence activity (Barbara Valent, personal communication). To identify amino acid changes responsible for loss of avirulence activity, avirulent and virulent alleles were compared (Fig. 2). The MgMEP1 genes in TH3 and Guy11 uniformly differed at four amino acid positions from all the avirulence conferring MgMEP1 and MgMEP2 genes: D82N, R87K, V173I, and Y191C relative to the O-137 MgMEP1 gene (Fig. 2). Site-directed mutagenesis analysis showed, however, that three of these changes (D82N, R87K, and Y191C) outside the peptidase motif had no or only a minor effect on avirulence function (Barbara Valent, personal communication). Conservative substitutions within the peptidase motif that did not change the defined peptidase motif (e.g., V173I, V173A, V173L, and V173T), eliminated avirulence activity. These results strongly suggest that any substitutions of V173 result in loss of avirulence activity. Among genes conferring avirulence, the G-213 MgMEP2 gene has the most divergent sequence when compared to the O-137 MgMEP1 gene. It contains 35 amino acid substitutions and one amino acid insertion, one of which is in the peptidase motif, V179L, a change which fit within the defined motif. Despite the high degree of sequence divergence, functional conservation observed in the G-213 MgMEP2 gene indicates that
certain sites, such as V173, appear critical for avirulence activity and have been conserved.

Figure 2. Comparison of avirulent and virulent MgMEP genes. MgMEP1 from O-137, G-1, G-78, and G-223, and MgMEP2 from G-213 confer avirulence (AVR), whereas MgMEP1 from TH3, Guy11 and 70-15 (indicated as 70-15a), and MgMEP3 from G-22 and 70-15 (indicated as 70-15b) lack avirulence function (avr). Asterisks (*) below the amino acids indicate the peptidase motif that is shared by the majority of zinc-dependent metallopeptidase [GSTALIVN]-x(2)-H-E-[LIVMFYW]-[DEHRK]-H-x-[LIVMFYWSPQ], in which [] indicates any one of the listed sequences is allowed, x indicates any sequence is allowed, and {} indicates any one of the listed sequences is not allowed. 2H’s and E in boldface are zinc ligands and the active site residue, respectively (http://ca.expasy.org/cgi-bin/prosite-search-ac?PDOC00129). In the text, amino acid residues are referred to by their position in the intact O-137 MgMEP1.
The G-22 \textit{MgMEP3} gene, which lacks avirulence activity, has a V173T substitution, in addition to numerous amino acids changes compared to gene family members conferring avirulence. It has other changes within the peptidase motif, but these substituted amino acids are allowed changes in the defined pattern.

2.3.4 \textit{MgMEP3} alleles lacked avirulence activity due to producing a truncated protein or a protein that is intact but probably fails to interact with \textit{Pi-ta}

The G-213 \textit{MgMEP2} gene conferred avirulence when introduced into the virulent strain CP987, but G-17 and G-22 \textit{MgMEP3} genes lacked avirulence activity. The G-17 \textit{MgMEP3} gene contains a frameshift mutation, resulting in a truncated protein (only 18 amino acids) if translated. Similar to the G-17 \textit{MgMEP3} gene, other isolates also carry a nonfunctional \textit{MgMEP3} gene because of a frameshift mutation (G-165, G-219, and G-220) or a nonsense mutation (G-221).

The G-22 \textit{MgMEP3} gene was predicted to produce an intact protein (226 amino acids), but failed to confer avirulence toward \textit{Pi-ta} rice. To determine if the lack of avirulence activity of the G-22 \textit{MgMEP3} is due to improper expression of the gene and/or the protein that may not interact with \textit{Pi-ta}, the functionality of the promoter and coding sequence (CDS) of the G-22 \textit{MgMEP3} gene was tested using chimeric genes. The promoter of the G-22 \textit{MgMEP3} was fused to the CDS of the O-137 \textit{MgMEP1} and G-213 \textit{MgMEP2} to generate \textit{P}_{\text{MgMEP3}} ::\text{C}_{\text{MgMEP1}} and \textit{P}_{\text{MgMEP3}} ::\text{C}_{\text{MgMEP2}}$, respectively, and the CDS of the G-22 \textit{MgMEP3} was fused to the promoters of the O-137 \textit{MgMEP1} and G-213 \textit{MgMEP2} to generate \textit{P}_{\text{MgMEP1}} ::\text{C}_{\text{MgMEP3}} and \textit{P}_{\text{MgMEP2}} ::\text{C}_{\text{MgMEP3}}$, respectively. Each of the chimeric genes was transformed into virulent \textit{M. grisea} strain CP987 using the \textit{Agrobacterium tumefaciens}-mediated transformation method described in Appendix. At least five transformants for each chimeric gene were confirmed by PCR to carry an intact introduced chimeric gene and then were assayed for resistance response in \textit{Pi-ta} rice. All the transformants carrying either \textit{P}_{\text{MgMEP1}} ::\text{C}_{\text{MgMEP3}} or \textit{P}_{\text{MgMEP2}} ::\text{C}_{\text{MgMEP3}} caused blast disease on \textit{Pi-ta} rice, suggesting that the CDS of the G-22 \textit{MgMEP3} encodes a protein
that does not trigger \textit{Pi-ta}-mediated resistance. Interestingly, two transformants carrying the $P_{MgMEP3}::C_{MgMEP1}$ and five transformants carrying the $P_{MgMEP3}::C_{MgMEP2}$ caused attenuated resistant responses in \textit{Pi-ta} rice but the rest of the transformants caused blast disease on \textit{Pi-ta} rice. This result suggests that the promoter of G-22 \textit{MgMEP3} might be functional although it may not be as active as those of \textit{MgMEP1} and \textit{MgMEP2}.

2.3.5 Phylogenetic analysis

To test if the occurrence of \textit{MgMEP} genes is specific to certain \textit{M. grisea} lineages, phylogenetic relationships of the \textit{M. grisea} isolates included in this study were determined. An organismal tree was inferred using the DNA sequences of the internal transcribed spacer (ITS) region of 23 \textit{M. grisea} isolates (Fig. 1B). The tree resolved three distinct lineages consisting of: (1) isolates from rice and a few grasses; (2) isolates from \textit{Digitaria} spp., and (3) isolates from \textit{Cyperus} spp. These relationships were supported with 100\% bootstrap values and also consistent with previous studies (BORROMEO \textit{et al.} 1993; COUCH and KOHN 2002; SHULL and HAMER 1994). Isolates from rice were further grouped (63\% bootstrap support) along with one isolate from rice cut-grass (\textit{Leersia hexandra}) which grows in rice fields as a weed. Three isolates from \textit{Pennisetum} spp. or \textit{Leersia oryzoides} were not grouped in any of lineages (G-78, G-223, and G-194). It was evident that the \textit{MgMEP1} and \textit{MgMEP2} genes were present in two distinct lineages, while the \textit{MgMEP3} gene was present only among rice isolates and the closely related isolates.

The gene tree showed that the \textit{MgMEP} gene family was divided into three clades (\textit{MgMEP1}, \textit{MgMEP2}, and \textit{MgMEP3}), which were supported by 100\% bootstrap values (Fig. 3). The division of the \textit{MgMEP} gene family was also supported by the significant DNA sequence identity between members in the same clade (96\% $\sim$ 100\%) and sequence divergence between members from different clades (69\% $\sim$ 93\%).
Figure 3. Gene genealogy of 23 *MgMEP* genes. A neighbor-joining tree was inferred from genomic coding sequences using the program MEGA2 Kimura 2-parameter model with complete deletion of gaps and both transitions and transversions substitutions included. Bootstrap values >50% of 1000 replicates are indicated near the nodes. The *MgMEP1*, *MgMEP2*, and *MgMEP3* are indicated by a filled rectangle, triangle, and circle, respectively, next to the names of the *M. grisea* isolate from which the gene was cloned. Identical sequences from multiple isolates are treated as a single taxon (e.g., G-1 / G-156 / G-189, G-160 / G-213 / G-32, and G-165 / G-219). A shaded circle indicates a pseudogene resulting from a frameshift mutation (e.g., a deletion at the position of 11th in G-17 or of 289th in G-165, G-219, and G-220, relative to the *MgMEP3* consensus sequence) or a nonsense mutation (e.g., C\(\rightarrow\)T at the position of 366th in G-221).
2.4 Discussion

The AVR gene MgMEP1 (previously called AVR-Pita) was originally cloned from rice isolate O-137 based on its avirulence function in that the gene prevents the fungus from infecting the rice cultivar containing the corresponding R gene Pi-ta (ORBACH et al. 2000). This study has shown that the MgMEP1 gene is a member of a new gene family that is widely distributed among isolates of M. grisea from diverse hosts (Fig.1). Two other family members, termed MgMEP2 and MgMEP3, were identified based on their homology to the O-137 MgMEP1 gene.

M. grisea can be differentiated into several genetically distinct host specific groups. In general, M. grisea isolates from cultivated crops (e.g., rice, wheat, common millet, foxtail millet, and finger millet) are genetically similar to one another, but distinct from strains isolated from Digitaria spp. (KATO et al. 2000). The strains used in this study were resolved into three distinct lineages (Fig. 1B): (i) isolates from rice and other grass species (ii) isolates from Digitaria spp., and (iii) isolates from Cyperus spp. These inferred phylogenetic relationships are consistent with those from previous studies (BORROMEO et al. 1993; COUCH and KOHN 2002; KATO et al. 2000). The MgMEP1 gene was widely distributed, including rice isolates and closely related strains as well as Digitaria isolates. Some of these genes identified in nonpathogens of rice were cloned and demonstrated to confer avirulence, indicating that they were functional homologs to the O-137 MgMEP1 gene. Another AVR gene, AVR1-CO39, which specifies avirulence to rice cultivar CO39, was cloned from a M. grisea isolate from weeping lovegrass (FARMAN et al. 2002). In addition, M. grisea isolate from finger millet was reported to have an AVR gene corresponding to the rice R gene Pi-a (YAEGASHI and ASAGA 1981). These results suggest that AVR genes that control rice cultivar specificity are common in nonpathogens of rice.
Significant difference in the degree of genetic stability was evident among *MgMEP* gene family members. The *MgMEP1* and *MgMEP2* genes appeared genetically unstable because substantial variation was observed with respect to the gene copy number in individual isolates (ranging from zero to three) and gene-linked restriction fragment length polymorphisms (RFLPs), while the *MgMEP3* gene appeared genetically more stable than the *MgMEP1* and *MgMEP2* genes because it did not exhibit much variation in its copy number and RFLPs (Fig.1A). The O-137 *MgMEP1* gene was extremely close to a telomere with its stop codon being separated from the telomere repeat sequence by 48 bp (Orbach *et al.* 1996). The telomeric O-137 *MgMEP1* appears to be genetically unstable because spontaneous mutants that gained virulence on Pi-ta rice cultivars frequently appeared under laboratory conditions (Valent and Chumley 1991). Molecular analysis of several such mutants identified different types of mutations responsible for the loss of avirulence. Some of them had a deletion, ranging in size from 100 bp to over 12.5 kb that spans either part of the *MgMEP1* gene or the whole gene plus its flanking sequences (Orbach *et al.* 2000). The hypervariability of these genes might be due to their telomeric location. Frequent deletions and rearrangements at chromosomal regions near the telomere have been observed in many organisms (Carlson *et al.* 1985; Charron and Michels 1988; Heather and Trask 2002). Thus, the dynamic nature of the *MgMEP1* and *MgMEP2* genes in many isolates may reflect their genomic location, as shown in the O-137 *MgMEP1* gene. Comparative analysis of genomic organization of these unstable *MgMEP1* and *MgMEP2* genes, and the stable *MgMEP3* gene may provide insight into the mechanisms that are responsible for virulence variation of this fungus.

The presence of both *MgMEP1* and *MgMEP3* genes in a haploid genome is a clear indication that they are paralogs derived from a gene duplication event. Interestingly, the *MgMEP3* gene was present only among rice isolates and their close relatives. This may suggest that the gene was duplicated only in those isolates after divergence from *Digitaria* isolates. The evolutionary significance of gene duplication is that it can lead to novel gene function and subsequent functional divergence of duplicated genes (Ohno 1970). Gene duplications appear to be prevalent because whole genome sequence
analyses of some eukaryotic genomes have shown that duplicate genes comprise perhaps 30 – 60% of the chromosomal genes (Ball and Cherry 2001).

Two identical copies of a gene will be produced by gene duplication. Once they are fixed in the population and preserved over time, these copies could evolve in three possible ways (Hurles 2004; Taylor and Raes 2004): (i) loss of function (nonfunctionalization) in that one copy becomes a pseudogene by accumulation of deleterious mutations; (ii) acquisition of new function (neofunctionalization) in that one copy retains its original function while the other accumulates mutations associated with new function; and (iii) splitting of functions (subfunctionalization) in that the two copies specialize to perform complementary functions by partitioning the original function between them. Analysis of the MgMEP3 gene revealed four interesting features: (i) lack of avirulence activity; (ii) a conserved peptidase motif; (iii) a functional promoter; and (iv) presence of pseudogenes due to a premature stop codon. The lack of avirulence and conserved peptidase motif may reflect the possible neofunctionalization of the MgMEP3 gene, which results in preservation of metabolic function while altering its recognition as an avirulence factor. The functionality of the MgMEP3 promoter, although it may be less active as or expressed differently from those of the MgMEP1 and MgMEP2 genes, indicates that it may have been also subjected to neofunctionalization in that the MgMEP3 gene has evolved to have a different spatiotemporal expression pattern. The human malarial parasite Plasmodium falciparum has two paralogous sets of ribosomal RNA genes that are expressed conditionally: S-type rRNA genes are expressed when the parasite is in the mosquito; A-type rRNA genes are expressed in the human host (Waters 1994). Further comparative analysis of the expression patterns between the MgMEP1 and MgMEP3 genes would test this hypothesis. The presence of a pseudogene is unique to the MgMEP member of the MgMEP family, although complete deletions of the MgMEP1 and MgMEP2 genes occurred in many isolates. The absence of a functional gene in some isolates suggests that the gene is unlikely to be essential. Alternatively, these genes may confer essential strain-specific functions. Positive evidence for these speculations may come from the revelation of the expression patterns and biological function of the MgMEP3 gene.
The metallopeptidase activity of the O-137 MgMEP1 gene appears critical for avirulence activity (ORBACH et al. 2000). All MgMEP genes cloned to date contain the conserved motif, suggesting that they are likely to encode functional peptidases. However, the presence of a conserved motif does not necessarily indicate the presence of avirulence activity. For example, conservative substitutions of the amino acid residue V173 that still fit the peptidase motif eliminate the virulence activity. This suggests that the exact structure of the active site of the peptidase is required for conferring avirulence activity. The presence of both avirulence and virulent natural alleles of the MgMEP genes in M. grisea population suggests that evolution of the MgMEP gene family involves modifications of a protein to preserve metabolic function while altering its recognition as an avirulence factor.
Chapter 3

COMPARATIVE ANALYSIS OF GENOMIC ORGANIZATION OF THE 
MGMEP GENE FAMILY AMONG ISOLATES OF MAGNAPORTHE GRISEA 
FROM DIVERSE HOSTS

3.1. Introduction

The haploid heterothallic ascomycete fungus Magnaporthe grisea (Herbert) Barr causes blast diseases on many gramineous species. This species can be differentiated into several genetically distinct host specific groups. In general, M. grisea strains isolated from cultivated crops (e.g., rice, wheat, common millet, foxtail millet, and finger millet) are genetically similar to one another, but distinct from strains isolated from Digitaria spp. (Kato et al. 2000). Rice pathogenic strains, within which many races have been identified according to their virulence or avirulence on particular rice cultivars (Zeigler et al. 1994), cause rice blast disease that limits rice production worldwide (Ou 1985). One of the desirable strategies for controlling the disease is the use of blast-resistant rice cultivars (Ou 1985; Zeigler et al. 1994). However, the resistance has not been durable because of the frequent appearance of new races with the ability to infect previously resistant cultivars. Potential mechanisms responsible for the evolution of new pathogenic races include heterokaryosis (Giatgong and Frederiksen 1969; Suzuki 1965), parasexual recombination (Genovesi and Magill 1976), and aneuploidy (Kameswar Row et al. 1985; Ou 1980). Because the interactions between M. grisea and rice follow a gene-for-gene model (Orbach et al. 2000; Zeigler et al. 1994), the mutations of avirulence (AVR) genes can lead to the generation of new races. Thus, understanding the molecular mechanisms underpinning the evolution and variation of AVR gene is critical to understanding how the pathogens overcome host resistance.

The AVR gene MgMEP1 (previously called AVR-Pita), which belongs to the MgMEP gene family, was cloned from M. grisea rice pathogenic field isolate O-137 using a map based cloning strategy (Orbach et al. 2000). The O-137 MgMEP1 prevents
infection of rice containing the corresponding resistance \((R)\) gene \(Pi-ta\). This gene was located extremely to a telomeric location on the chromosome with its stop codon separated from the telomere repeat sequence by 48 bp (Fig. 1A). The telomere repeat sequence of \(M. grisea\) consists of an array of hexanucleotide, TTAGGG (Farman and Leong 1995; Orbach et al. 2000) also found in the telomeres of vertebrates, slime molds and other fungi, including Neurospora crassa (Schechtman 1990), Cladosporium fulvum (Coleman et al. 1993), and Fusarium oxysporum (Powell and Kistler 1990).

The telomeric O-137 MgMEP1 gene appears to be unstable because spontaneous mutants that gained virulence on \(Pi-ta\) rice cultivars frequently appeared under laboratory conditions (Valent and Chumley 1991). Molecular analysis of several such mutants identified different types of mutations responsible for the loss of avirulence. Some of them had a deletion, ranging in size from 100 bp to over 12.5 kb, that spans either part of the MgMEP1 gene or the whole gene plus its flanking sequences (Orbach et al. 2000). Other mutants had a point mutation within the gene (Orbach et al. 2000). Another telomere-linked \(AVR\) gene in the same strain (O-137), \(AVRI-TSUY\), also mutates at a high frequency (Kang et al. 2000). The hypervariability of these genes might be due to their telomeric location, because frequent deletions and rearrangements at chromosomal regions near the telomere have been observed in many organisms (Carlson et al. 1985; Charron and Michels 1988; Heather and Trask 2002).

Repetitive DNA sequences also may contribute to pathogen variability. Repetitive DNA sequences can be classified into two major groups, including (i) tandemly repeated sequences such as ribosomal RNA genes and satellite DNAs, and (ii) dispersed repeats such as transposable elements. Transposable elements can be divided into two classes according to their mode of transposition: retrotransposons that transpose through reverse transcription of an RNA intermediate and DNA transposons that transpose directly by a cut-and-paste mechanism (Galun 2003). Retrotransposons can be further split into two subclasses, LTR retrotransposon and non-LTR retrotransposon, depending on the presence or absence of long terminal repeat (LTR) sequences at both ends. These repetitive elements can promote a series of genome rearrangements through recombination between homologous repeats located on the same or different
chromosomes. Transposition events of transposable elements often cause insertional mutations. For example, one spontaneous gain-of-virulence mutant on Pi-ta rice contained an insertion of a Pot3 transposon in the putative promoter region of the O-137 MgMEP1 gene, causing the disruption of gene expression (KANG et al. 2001). This study provides direct evidence that transposition activity can change host range in M. grisea.

Two other family members, termed MgMEP2 (functional as AVR gene) and MgMEP3 (nonfunctional as AVR gene), were cloned based on their homology to the O-137 MgMEP1 gene. A survey of M. grisea strains from diverse hosts through Southern blot analysis revealed two interesting features: (1) MgMEP1 and MgMEP2 were present widely even among strains that are genetically distant from the rice pathogens, but they appeared genetically unstable because substantial variation was observed with respect to the copy number of the genes in individual strains (ranging from zero to three) and gene-linked restriction fragment length polymorphisms (RFLPs); (2) MgMEP3 was present only among rice pathogens and those that are genetically close to rice pathogens. Relative to MgMEP1 and MgMEP2, this gene appeared to be genetically more stable because it did not exhibit much variation in copy number and RFLPs.

A better understanding of the mechanisms underpinning the evolution of the MgMEP gene family in M. grisea can be obtained by characterizing the genomic organization of the gene family. In this study, nine genomic regions carrying a MgMEP gene family member were cloned from eight different M. grisea strains and compared.

3.2. Materials and Methods

3.2.1. Strains and growth conditions

The M. grisea field isolates used in this study are described in Table1. Escherichia coli strain XL1-blue MRF’ was used for maintaining plasmids.
3.2.2. Cloning by genomic library screening

Clones containing the \textit{MgMEP1}, \textit{MgMEP2}, or \textit{MgMEP3} gene were isolated from genomic DNA libraries of non-rice infecting \textit{M. grisea} field isolates G-1, G-17, G-22, G-78, G-213, and G-223, constructed using the \textit{\lambda}GEM12 vector (Promega, Madison, WI). Genomic libraries were screened with a \textsuperscript{32}P-labeled O-137 \textit{MgMEP1} probe. Positive plaques were purified, and \textit{\lambda} DNAs isolated using a plate lysate method. Electrophoresis of restriction enzyme-digested lambda DNA and Southern blotting were performed according to standard protocols (Sambrook \textit{et al.} 1989). Restriction fragments hybridizing to the probe were subcloned into pGEM3Zf and sequenced by utilizing the EZ::TN transposon insertion system (Epicentre Technologies, Madison, WI) and/or primer walking.

3.2.3. Inverse PCR

Inverse polymerase chain reaction (PCR) is a method for amplifying DNA sequences that flank a region of known sequence. The method is based on PCR using a pair of primers that are oriented in the reverse direction and a self-ligated restriction fragment as a DNA template (Ochman \textit{et al.} 1988). The protocol used in this study was adapted from the original protocol (Ochman \textit{et al.} 1988). The DNA template was prepared by digesting 1 µg of G-1 genomic DNA with \textit{XmnI} at 37°C for 12 h in a total volume of 100 µl. Completely digested DNA was extracted with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) and the aqueous phase was recovered by centrifugation for 5 min at 4°C. The DNA was precipitated with 10 µl of 3 M Na-Acetate, pH 5.5 and 200 µl of 100% ethanol, followed by 30 min incubation at −20°C. The DNA was pelleted by centrifugation, air-dried, and resuspended in 50 µl TE buffer (10mM Tris-HCl, 1 mM EDTA, pH8.0). The DNA was self-ligated for 16 h at 16°C in a 500 µl of reaction volume. The ligation sample was ethanol precipitated as described above and resuspended in 30 µl TE buffer. Inverse PCR was performed in a 20 µl reaction mixture containing 10 µl of template DNA and 2 µl (10 pmol/µl) each of the following two
primers: pita-1 (5’-TCATGCTCCCTGACTTTTGA-3’) and pita-2 (5’-GCCGCGCTAAAAAAGGTAAAT-3’), using the FailSafe PCR system (Epicentre Technologies, Madison, WI). The PCR cycling program consisted of an initial denaturation for 5 min at 95°C, 30 cycles of 30 sec denaturation at 95°C, 30 sec annealing at 60°C, 2 min extension at 72°C, and followed by a final extension for 10 min at 72°C. PCR products were cloned into a pGEM-T Easy vector (Promega, Madison, WI). Three individual PCR clones were included in DNA sequence analysis to identify potential sequence changes introduced during PCR amplification.

3.2.4. DNA sequencing and computer analysis

DNA sequencing reactions were performed using the ABI Prism Big-Dye terminator sequencing kit (Applied Biosystems, Foster City, CA) and analyzed with an ABI377 sequencer. DNA reads were edited manually for ambiguous sites and assembled using Autoassembler (Applied Biosystems, Foster City, CA). The dotplot program in the Lasergene package (DNASTar, Madison, WI) was used to identify repeats within a sequence and to compare two sequences for the presence of homologous regions. The computer program MEGA2 (Kumar et al. 2001) was used to compute the frequencies of transition/transversion.

3.2.5. Database searching

The GenBank/EMBL/DDBJ (http://www.ncbi.nlm.nih.gov/BLAST/) databases were searched to identify putative genes in the sequences flanking the MgMEP gene family using BLAST. The Magnaporthe grisea genome sequence database (http://www.broad.mit.edu/annotation/fungi/magnaporthe/index.html) was searched to identify contigs and supercontigs containing MgMEP gene family members and transposable elements using BLAST. The MagnaportheDB database
(http://www.fungalgenomics.ncsu.edu/Projects/mgdatabase/int.htm) was used to locate the genes on physical and genetic maps.

3.3 Results

3.3.1 Some members of the \textit{MgMEP} gene family are closely linked to the telomere

The chromosomal locations of the \textit{MgMEP1} and \textit{MgMEP2} genes were determined in \textit{M. grisea} field isolates G-1 and G-213, isolated from \textit{Digitaria} spp., and G-78 and G-223, isolated from \textit{Pennisetum} spp (Table 1). Similar to the O-137 \textit{MgMEP1} gene, the \textit{MgMEP1} gene in isolate G-1 was found to be telomeric (Fig. 4A). The telomeric location of the G-1 \textit{MgMEP1} gene was revealed by chromosome-walking from the G-1 \textit{MgMEP1} gene by inverse PCR (OCHMAN et al. 1988) to amplify a DNA fragment containing the upstream region of the gene. A 1.7 kb amplicon was obtained and subsequently cloned into pGEM-T Easy vector, three clones from which were sequenced. One end of all three clones had the same sequences that overlapped with the 5’ flanking region of the \textit{MgMEP1} gene, indicating that inverse PCR successfully amplified sequences flanking the upstream region of the gene. The other ends of the clones had distinctive telomere repeat sequences. Two of the clones contained 29 complete repeats of (TTAGGG) and one incomplete repeat of (TTAGG). The third clone contained 20 complete repeats and one partial repeat of (TTAGG). The differences in the numbers of telomere repeat in three PCR clones may reflect heterogeneity among telomeric repeat sequences at this chromosome in different nuclei. However, the possibility of experimental artifacts (such as a truncation of the repeats during genomic DNA preparation, PCR, and/or propagation in \textit{E. coli}) cannot be ruled out. Successful ligation of telomere repeats to the blunt ended \textit{XmnI}-digested genomic DNA suggests that the chromosome end where the \textit{MgMEP1} gene is located has a blunt end.
Although both the O-137 MgMEP1 and G-1 MgMEP1 genes were found located near a telomere, their transcriptional orientations relative to the telomeric repeats were opposite (Fig. 4A). The stop codon of the O-137 MgMEP1 gene is separated from the telomere repeat by 48 bp, whereas the start codon of the G-1 MgMEP1 gene is 1889 bp away from the telomere repeat.

Strains O-137 and G-1 were isolated from rice (Oryzae sativa) and Digitaria spp., respectively. These two strains have been recognized as members of two separate phylogenetic species (COUCH and KOHN 2002; KATO et al. 2000). Given the telomeric location of MgMEP1 in two isolates from different species the MgMEP1 gene in other strains was hypothesized to be telomeric. To test the hypothesis, a Southern analysis was performed using the MgMEP1 gene and the telomeric repeat sequences as probes (S-Y Park, personal communication). The identification of a fragment that hybridizes to both probes and the size of the fragment would indicate telomeric location of the gene and provide an estimate of physical distance between the gene and a telomere. The MgMEP1 genes from the strains G-223 and G-78 were found to be telomeric, with distances from the repeat ranging from 10 kb to 30 kb (S-Y Park, personal communication). Using the same approach, the MgMEP2 gene from G-213 was also found to be telomeric (S-Y Park, personal communication). The corresponding fragments have been cloned from genomic libraries from each isolate enriched for telomeric fragments and sequencing analysis is in progress (S-Y Park, personal communication).
Figure 4. Comparison of the genomic organizations of the *MgMEP* loci in different *M. grisea* isolates. O-137 sequence (6214 bp) was given by Dr. Valent, and 70-15 sequences (contigs 2.1975, 2.1976, and 2.2232) were retrieved from the *M. grisea* genome database (http://www.broad.mit.edu/). The other sequences were determined in our laboratory (see Materials and Methods). A putative coding sequence is indicated by an open box with a corresponding gene name. A pseudogene is indicated by a ψ, followed by a gene name. The telomere repeat sequence is indicated by the box with T. Internal domains of novel retrotransposons (e.g., MGLR-4 and MGLR-5), their associated LTR, and solo-LTR are marked with boxes filled with different patterns or colors. The orientation (5’ to 3’) of transcription is indicated by an arrow above or inside a box. A, The *MgMEP1* and *MgMEP2* loci were compared among six isolates. A broken vertical line indicates conserved regions immediately upstream and downstream of the *MgMEP* coding sequences. B, The *MgMEP3* locus was compared among three different isolates. The corresponding conserved regions are indicated by a double-line and broken vertical line.
Previously, it had been suggested that \textit{MgMEP1} and \textit{MgMEP2} genes are allelic because no strains have been found to carry both genes, although \textit{MgMEP1} and \textit{MgMEP3} appeared together in several strains. In addition, the G-213 \textit{MgMEP2} gene has significant DNA sequence identity (92%) with the O-137 \textit{MgMEP1} gene and fully complemented the function of the O-137 \textit{MgMEP1} gene as avirulence gene. Similar chromosomal locations of these genes further support their allelic relationship.

The chromosomal locations of the \textit{MgMEP3} gene present in field isolates G-17 and G-22 have not been determined directly. However, the regions containing the \textit{MgMEP3} gene from these strains had a strong syntenic relationship with the \textit{MgMEP3} gene in 70-15, which is distant from the telomere, suggesting that the \textit{MgMEP3} gene is unlikely to be telomeric (see below).

3.3.2 Genomic location of the \textit{MgMEP1} and \textit{MgMEP3} genes in \textit{M. grisea} strain 70-15

The \textit{M. grisea} strain 70-15 is a laboratory strain that has been chosen for genome sequencing. 70-15 was derived from a cross between the rice infecting strain Guy 11 and a weeping lovegrass infecting strain, followed by numerous back crosses to Guy 11 (CHAO and ELLINGBOE 1991; LEUNG et al. 1988). \textit{M. grisea} contains a haploid genome consisting of 7 nuclear chromosomes totally ~40 Mb (ORBACH et al. 1996; TALBOT et al. 1993). A high-density genetic map has been constructed to establish seven linkage groups (NITTA et al. 1997). A current draft version of the genome assembly consists of 2273 contiguous stretches of sequence (contigs; 17 kb in average length) that are ordered and linked together into 159 supercontigs (243 kb in average length). Currently, 37 supercontigs, representing 35.6 Mb (93% of the assembly) have been ordered on seven linkage groups that correspond to the seven chromosomes (MARTIN et al. 2002).

Both \textit{MgMEP1} (locus ID MG11081.4) and \textit{MgMEP3} (no locus ID assigned) genes exist in the 70-15 genome sequence. The \textit{MgMEP1} gene was found in the contig 2.2232 (3251 bp) in supercontig 124 (5281 bp). The chromosomal location of supercontig 124 remains to be determined. The 70-15 \textit{MgMEP3} gene was found in contig 2.1976 (5787 bp) in supercontig 32 (259,942 kb), known to be located on linkage group VII
(chromosome 7). Although the precise physical distance between the *MgMEP3* gene and the telomere remains to be estimated, the *MgMEP3* gene is separated from the telomere by at least 250 kb because another supercontig 33 (250,947 kb), which was connected to supercontig 32, is located at one end of the linkage group.

### 3.3.3 Identification and characterization of two novel gypsy-type retrotransposons that are closely associated with *MgMEP* genes

LTR retrotransposons can be subdivided into two major types, the Ty1/copia-type and the gypsy/Ty3-type (or simply gypsy-type), according to the arrangement of their protein-coding internal domains (Galun 2003). Fungal gypsy-type retrotransposons are 5 – 8 kb long and have 200 bp- 600 bp long LTRs flanking the internal genes encoding proteins similar to the gal and pol retroviral gene products. The *pol* genes encode all proteins required for transposition, including protease (PR), reverse transcriptase (RT), RnaseH (RH) and integrase (IN) (Daboussi 1996; Galun 2003). Unequal crossing over and/or intrachromosomal recombination between the LTRs can remove the internal genes, generating a solo LRT (sequences containing only the LTR DNAs) at the recombination site (Bennetzen and Kellogg 1997). Four gypsy-type retrotransposons found in *M. grisea* to date include grasshopper (Robinson et al. 1993), MAGGY (for Magnaporthe grisea gypsy-type) (Farman et al. 1996), MGRL-3 (*M. grisea* LTR retrotransposon-3) (Kang 2001), and Pyret (Nakayashiki et al. 2001). Sequence analysis of regions flanking the *MgMEP1* and *MgMEP2* genes identified DNA repeat sequences that include a REP1 element (Farman et al. 2002), MGR608 and MGR609 (Hamer et al. 1989; Kang et al. 1995). Previously, the REP1 element (188 bp) was identified as a repetitive sequence found at the deletion borders of the *AVR-CO39* gene in certain *M. grisea* field isolates (Farman et al. 2002). MGR608 (126 bp) and MGR619 (79 bp), which belong to a family of dispersed, repeated DNA sequences called MGR (for Magnaporthe grisea repeat) (Hamer et al. 1989), were frequently found together as a contiguous sequence, thereafter referred to as MGR619/MGR608 (205 bp).
MGR619/MGR608 was closely associated with the PWL gene family (Kang et al. 1995) as well as the MgMEP gene family.

The REP1 and MGR619/MGR608 sequences flanking MgMEP genes existed in many deleted or fused forms with other uncharacterized repetitive sequences. Dispersed repetitive sequences have been speculated as deletion derivatives or partial sequences of unidentified transposable elements (Hamer et al. 1989). Interestingly, neighboring sequences of some of these repeats showed significant similarity to the gypsy-type retrotransposons. Subsequently, it was hypothesized that the REP1 and MGR619/MGR608 were the LTRs of putative retrotransposons. Lack of sequence similarity between REP1 and MGR619/MGR608 indicates that they may belong to two different retrotransposons. Because there was no complete copy of retrotransposons in areas flanking the MgMEP genes, a reference copy was first identified by searching the genome sequences of strain 70-15. BLASTN searches of the M. grisea genome database using either REP1 or MGR619/MGR608 as a query identified 77 and 100 contigs, respectively, indicating that they were prevalent in the genome. Some of the contigs contained multiple copies of these sequences.

One contig containing REP1, 2.1803, harbored two annotated genes, MG09406 and MG09407. MG09406 (879 bp) encodes a putative gag protein (292 amino acids) with a putative zinc-finger domain (Cys-X_2-Cys-X_9-Cys) found in various retrotransposons. The second gene MG0907 (2995 bp), which starts 485 bp downstream from the stop codon of MG09406, encodes a putative pol protein (755 amino acids). These two genes were flanked by direct repeats (193 bp and 194 bp; one insertion/deletion found). These direct repeats contain the 5’ and 3’ terminal sequences (5’ TG…CA 3’) found in most retrotransposon LTRs (Galun 2003). However, they lacked inverted terminal repeats present in most LTR-type retrotransposons, including grasshopper and MAGGY (Dobinson et al. 1993; Farman et al. 1996). These LTRs were flanked by a 5 bp direct repeat, ATGAG, which is characteristic of a target site duplication generated during transposition. This novel retrotransposon, termed MGRL-4 (Magnaporthe grisea LTR retrotransposon-4), exhibited significant similarity to MGRL-3 at the reverse transcriptase domain previously used to study the evolutionary relationship among
various retrotransposons and viruses (Capy et al. 1996; Xiong and Eickbush 1990). Structural organization of the four enzymes (in the order of PR, RT, RH, and IN) encoded by the *pol* gene suggests that MGLR-4 is a gypsy-type retrotransposon like MGLR-3 (Kang 2001).

The REP1 (158 bp) sequence associated with the *MgMEP* genes was identical to the corresponding region of the LTRs (193 bp and 194 bp) of MGLR-4, indicating that REP1 was part of MGLR-4. Sequence comparison confirmed that the putative retrotransposon found in strain G-1 located adjacent to REP1 was indeed a truncated copy of MGLR-4. This truncated copy existed between the telomere and the upstream region of the G-1 *MgMEP1* gene (Fig. 4A). When compared to the reference copy of MGRL-4 (5875 bp), the truncated copy had only a 3’ region (1181 bp), immediately followed by telomere repeat sequences. Its 3’ LTR was flanked by a partial solo LTR sequence derived from another retrotransposon. In addition to the deletion of its 5’ region, the predicted protein sequences of the remaining part were interrupted several stop codons. Comparison of the sequences between the truncated MGRL-4 and the corresponding region in an intact copy of MGLR-4 revealed 95% (1122/1181 nucleotides) identity. Interestingly, most of the differences (96%, 55/57 nucleotides; two deletions excluded) were due to C:G to T:A transitions, characteristic results of Repeat-Induced Point mutation (RIP) or RIP-like processes that inactivate repeated sequences during the sexual cycle (Cambareri et al. 1989; Selker et al. 1987).

Similar to the REP1 sequence, MGR619/MGR608 was part of an LTR of another novel retrotransposon, termed MGLR-5. An intact copy of MGLR-5 was found in the contig 2.366, which was identified by using BLAST against the genome sequence with MGR619/MGR608 as a query. The matching contig contained only one annotated gene MG01939.4 (5502 bp) encoding a 1689 amino acid protein exhibiting significant similarity to gag and pol proteins (consisting of four enzymes in the order of PR, RT, RH, and IN). This suggests that MGLR-5 is also a gypsy-type retrotransposon. The gene was flanked by perfect direct repeats (208 bp of LTRs) that exhibited significant similarity (88%, 184/208 nucleotides) to MGR619/MGR608 (205 bp), suggesting that MGR619/MGR608 is a derivative of the LTR of MGLR-5. Sequence comparison
confirmed that the putative retrotransposons found adjacent to the MGR619/MGR608 element near the \textit{MgMEP} genes in several strains (e.g., G-1, G-223, G-78, and G-213) were indeed truncated copies of MGLR-5. The G-213 \textit{MgMEP2} gene was flanked by two identical copies of MGLR-5 in a head-to-head orientation. The orientation of MGLR-5 in the 3’ region of the G-1 and G-223 \textit{MgMEP1} gene was opposite to that in the G-78 and G-213. Numerous stop codons in their internal domains indicate that they are degenerate copies.

Pairwise comparison of all the MGLR-5 copies including an intact copy revealed that most polymorphisms resulted from C:G to T:A transitions, again suggesting that RIP-like mechanisms are responsible for most of their nucleotide diversity. The distribution pattern of MGLR-5 within various host-specific forms of \textit{M. grisea} was investigated (Fig. 5). \textit{EcoRI}-digested genomic DNA of 27 strains from diverse hosts were hybridized with a 320 bp \textit{Hpal-BamHI} fragment of MGLR-5 corresponding the 3’ end of the internal domain (Fig. 5). All 27 strains (Table 1) contained sequences homologous to MGLR-5. This probe hybridized to many bands in some isolates, indicating that MGLR-5 was present in high copy numbers in the genomes of these isolates. The strains G-213 and G-32 were isolated from the \textit{Digitaria} spp. in Japan and carried an identical copy of \textit{MgMEP2}. Southern blot analysis showed that these two strains had different MGLR-5 profiles, with a 3.6 kb band that exhibited a strong signal relative to other bands in these strains. When digested with \textit{EcoRI}, G-213 genomic DNA releases a 3.6 kb fragment that contains two identical copies of MGLR-5 flanking the \textit{MgMEP2} gene. Because this fragment has two binding sites for the probe, the hybridization signal was likely to be higher in this fragment. The presence of the same band in the G-32 indicates that the G-32 \textit{MgMEP2} gene also may be flanked by two copies of MGLR-5.
Figure 5. Restriction map and organization of the G-213 MgMEP2 locus and Southern blot analysis of genomic DNA from 27 Magnaporthe grisea field strains. A. Restriction map of 7.1 kb of the MgMEP2 locus in the strain G-213 that carries two identical copies of MGLR-5 in a head-to-head orientation. Solo LTRs are marked as in Fig. 1. The orientation (5’ to 3’) of transcription is indicated by an arrow above or inside a box. Selected restriction sites are marked as follows: H, HindIII; E, EcoRI; Hp, HpaI and B, BamHI. The positions of the hybridization probe Hp-B are shown at the bottom of the map. B. Genomic DNA was digested with EcoRI and hybridized with Hp-B probe (320 bp). The genus (or common name) of the host plant from which individual strains (Table 1) were isolated is shown on top of each panel, where known. An arrow head indicates the band expected in the G-213 that had two probe binding sites because of the presence of two copies of MGLR-5.
Different forms of solo LTRs originating from MGLR-5, MGLR-6 and Pyret were found in the region flanking the \textit{MgMEP} gene family (Fig. 4). In some strains (O-137, G-1, G-223, and 70-15), the solo LTR derived from MGLR-5 was located upstream of the \textit{MgMEP1} gene at the same position. Interestingly, this solo LTR was truncated by MGLR-4 in G-1 or by another solo LTR derived from MGLR-5 in G-223 and 70-15. In the 3’ region of the 70-15 \textit{MgMEP1} gene, three different solo LTRs were found.

3.3.4 Syntenic relationship in the regions containing the \textit{MgMEP1} and \textit{MgMEP2} genes

Genome sequences at the \textit{MgMEP1}/\textit{MgMEP2} locus were compared among six strains comprising five field strains (O-137, G-1, G-223, G-78, and G-213) and one laboratory strain 70-15. For O-137, a telomeric 6214 bp region (ORBACH \textit{et al.} 2000) was used. For G-1, G-223, G-78, and G-213, subclones (4486 bp, 5710 bp, 5013 bp, and 7779 bp, respectively) were obtained from each genomic \textit{\lambda} clone and sequenced. For the laboratory strain 70-15, contig 2.2232 (3251 bp), containing the \textit{MgMEP1} gene, was retrieved from the genome database. As described above, all the \textit{MgMEP1} and \textit{MgMEP2} genes cloned to date were flanked by retrotransposons and solo LTRs, whose locations and orientations relative to the \textit{MgMEP} genes varied among \textit{M. grisea} strains. Significant sequence similarity existed only at the putative promoter region of the \textit{MgMEP} genes and the 3’-untranslated region, which were interrupted by LTRs. Additional sequences beyond the LTRs were available only for some strains including O-137, G-223, 70-15, and G-78. Sequence comparison found no syntenic relationships among these sequences.

3.3.5 Syntenic relationship in the regions containing the \textit{MgMEP3} gene in G-17, G-22, and 70-15

Genomic sequences of the \textit{MgMEP3} locus among three different strains (G-17, G-22, and 70-15) were compared. For G-17 and G-22, subclones (2575 bp and 8134 bp, respectively) were obtained from each genomic \textit{\lambda} clone and sequenced. For the laboratory
strain 70-15, contig 2.1976 (5787 bp), which contains \textit{MgMEP3}, was retrieved from the genome database as well as its neighboring contig 2.1975 (9945 bp). Since the genome sequences of 70-15 are draft sequences, it is not certain whether these two contigs together form an ungapped contig. In this study, however, they were considered as one contiguous sequence. The conserved regions (more than 93% identity) between two sequences were indicated by a vertical broken line (Fig. 4B). Although the sequence of the \textit{MgMEP3} locus appeared to be conserved among three isolates G-17, G-22, and 70-15, which probably is due to the close relationship of these strains (Fig. 4B), the syntenic relationship between strains was interrupted by large insertions (or deletions). For example, the \textit{MgMEP3} locus of G-17 (2575 bp) exhibited 97% identity to the corresponding region of G-22. Downstream of the gene, however, a fragment of 198 bp present in G-22 was missing in both G-17 and 70-15. No DNA repeat sequences responsible for the deletion were apparent in this missing sequence or flanking sequences. Upstream of the gene, an additional sequence block (329 bp) existed only in 70-15, which is likely to be part of the putative promoter considering its vicinity to the \textit{MgMEP3} gene. It remains to be determined if this sequence is involved in regulation of gene expression. Disruption of the synteny became apparent at the regions further upstream between G-22 and 70-15, due to the insertions of an unknown transposon and other repeat sequences (e.g., Pot3 and solo-LTR derived from Pyret).

3.3.6 Syntenic relationship between the regions of the \textit{MgMEP1/MgMEP2} genes and \textit{MgMEP3} genes

The genomic coding sequences of the O-137 \textit{MgMEP1} and G-22 \textit{MgMEP3} genes exhibit 72% DNA sequence identity. When the flanking sequences of the O-137 \textit{MgMEP1} and G-22 \textit{MgMEP3} genes were compared, significant sequence identity (65%) was found only at the region 180 bp upstream from the translation start site, and no similarity was found immediately after the stop codon. The upstream region of the 70-15 \textit{MgMEP3} gene had solo LTRs that derived from MGLR-5 and Pyret, also found in the 70-15 \textit{MgMEP1} locus.
3.4. Discussion

This study shows that the \textit{MgMEP1} and \textit{MgMEP2} genes in some field isolates are located in the vicinity of telomeres, at distances ranging from 48 bp to 30 kb. In O-137 and G-1, the transcriptional orientations of the \textit{MgMEP1} gene relative to the telomere were opposite, which may have resulted from an inversion event of the DNA segment containing the gene. The \textit{M. grisea} strains included in this study, except the laboratory strain 70-15, were collected from diverse host species including those from \textit{Digitaria} spp. (G-1 and G-213) described as members of a separate species from isolates from rice (Couch and Kohn 2002; Kato et al. 2000). Therefore, variation in the distance between the telomere and the gene, and the inverted gene orientation found in O-137 and G-1 may reflect the dynamic evolution of the telomeric \textit{MgMEP1/MgMEP2} genes preceding speciation into several genetically distinct host specific groups.

Telomeres generally consist of a track of simple repeats and a set of associated proteins. Telomeres are essential structures because they protect the chromosome ends from end-to-end fusion or exonucleolytic degradation (Blackburn 1991). Subtelomeres correspond to the regions between the telomere repeat sequence and chromosome-specific sequences (Heather and Trask 2002). Subtelomeres structurally vary greatly among organisms, but typically contain various types of DNA repeats and pseudogenes (Heather and Trask 2002). The subtelomere is highly dynamic in that frequent deletions and rearrangements of the DNA sequences have been observed in many organisms (Carlson et al. 1985; Charron and Michels 1988; Heather and Trask 2002). The highly dynamic nature of the genes located in the subtelomeric region has been illustrated by the \textit{TLH} (telomere-linked helicase 1) gene family in \textit{M. grisea}. Most family members are located within 10 kb from telomere repeat sequence and undergo frequent deletion and amplification events (Gao et al. 2002).

Strain O-137 produced spontaneous mutants that had different lengths of deletion of the telomeric \textit{MgMEP1} gene (Orbach et al. 2000). The \textit{MgMEP1/MgMEP2} genes from \textit{M. grisea} strains isolated from diverse hosts also exhibited genetic variation in their location and orientation relative to the telomere repeat sequences. These observations
suggest that genetic instability appears to be a general feature of subtelomeres in *M. grisea* chromosomes. The dynamic nature of the subtelomere may result from frequent ectopic (non-allelic) recombination. In *Saccharomyces cerevisiae*, many chromosome ends contain several copies of the Y’ element that are tandemly arrayed adjacent to the telomere repeat sequences (LOUIS 1995; WALMSLEY *et al.* 1984). Frequent ectopic recombination between subtelomeres resulted in duplication or loss of Y’ elements on different chromosomes ends (LOUIS and HABER 1990; LOUIS and HABER 1992; LOUIS *et al.* 1994). The subtelomeres of the human malarial parasite *Plasmodium falciparum* also undergo frequent ectopic recombination. The physical clustering of subtelomeres at the *P. falciparum* nuclear periphery is mediated by subtelomeric homology, which is likely to facilitate ectopic recombination (FREITAS-JUNIOR *et al.* 2000). *P. falciparum* has antigenic var genes in the subtelomeric regions and their frequent recombination promotes variation that helps the parasite evade the host immune system (FREITAS-JUNIOR *et al.* 2000). *M. grisea* has several AVR genes, including AVR1-Ku86, AVR1-MedNoi, and PWL1, that are mapped to the telomere (DIOH *et al.* 2000; KANG *et al.* 1995). Given the dynamic nature of the subtelomeres, the presence of AVR genes in these regions may provide an advantage to the fungus by allowing it to rapidly overcome newly deployed R genes.

The most notable feature in the regions flanking the MgMEP1/MgMEP2 genes was the presence of diverse repetitive DNA sequences. They include two novel retrotransposons, MGLR-4 and MGLR-5, solo LTRs, and two inverted terminal repeat (ITR) transposons with significant similarity to Pot2 (KACHROO *et al.* 1994) and Pot3 (KANG *et al.* 2001). DNA repeat sequences have been implicated as a major source of genetic variation. Transposition events often cause loss-of-function mutations that can change the virulence spectrum of pathogens (KANG *et al.* 2001). The repetitive nature of these elements can promote a series of genome rearrangements through recombination between dispersed repeats located on the same or different chromosomes, resulting in deletion, duplication, inversion, and translocation, depending on the relative orientation and position of combining repeats.
In *M. grisea*, chromosomal rearrangements often are associated with clusters of transposable elements, suggesting that they result from recombination between transposable elements present throughout the genome (Nitta et al. 1997). All characterized *PWL* genes are flanked by diverse DNA repeat sequences that appear to be responsible for the deletion/amplification of the gene family. Three tandem copies of the *PWL2* gene present in a field strain were likely to be generated through successive unequal crossing over between homologous DNA repeat sequences present on both sides of the *PWL2* gene (Kang et al. 2000). The solo LTRs are generated by unequal crossing over and/or intrachromosomal recombination between the LTRs of a retrotransposon that results in excision of the internal domains (Bennetzen and Kellogg 1997). Thus, the presence of solo LTRs reflects previous transposition and subsequent recombination events. The *MgMEP1/MgMEP2* genes are closely associated with solo LTRs, indicating active recombination events. Deletion and amplification of the genes observed in certain *M. grisea* strains is likely to have resulted from recombination between these repeat sequences. In contrast to the highly dynamic nature of the *MgMEP1/MgMEP2* genes, the *MgMEP3* gene appeared genetically stable. The relatively low number of repeats in this gene may explain its stability in *M. grisea* populations.

Another possible role of transposable elements is to modulate the expression of neighboring genes. Since the discovery that LTRs carried promoter and enhancer motifs (Sverdlov 1998), many examples have shown that LTRs and parts of the retrotransposons influence the expression of adjacent genes (Brosius 1999). The O-137 *MgMEP1* gene has three solo LTRs in its upstream regions (Fig. 4A). A deletion analysis of the putative promoter region of the O-137 *MgMEP1* gene suggests that the solo LTRs are not required for expression of the gene. However, the possibility that these solo LTRs play a role in spatio-temporal regulation, which has been observed in other organisms (Brosius 1999), remains to be tested.

Considering the highly dynamic nature of the *MgMEP1/MgMEP2* genes and their avirulence functionality that potentially limits host range, these genes would have disappeared in the *M. grisea* population if they did not confer fitness. A number of bacterial *AVR* genes have been shown to be necessary for the fitness of the pathogen
Some fungal AVR genes have also been demonstrated to contribute to pathogen fitness, including ECP1 and ECP2 (ExtraCellular Protein 1 and 2) from Cladosporium funvum (Laugé et al. 1997) and NIP1 from Rhynchosporium secalis (Rohe et al. 1995). Accumulating evidence indicates that the maintenance of AVR genes is due to their function as fitness factors. The O-137 MgMEP1 gene that was originally identified on the basis of its avirulence activity appears to have another function as a fitness factor, which is supported by possible metallopeptidase activity (B. Valent, personal communication) and infection-specific induction of gene expression (Orbach et al. 2000). The maintenance and conservation of MgMEP1 genes further supports their important role in pathogen fitness. However, the absence of the genes in some strains suggests that their function is unlikely to be essential. Alternatively, these genes may confer strain or lineage-specific function.

Previous phylogenetic analysis indicates that the MgMEP1/MgMEP2 and MgMEP3 genes are paralogous to one another and arose as a result of a gene duplication event (Chap. 2). The MgMEP1/MgMEP2 and MgMEP3 genes appear to have different chromosomal locations. Such dispersed paralogs can be generated through a retrotransposition mechanism involving the reverse transcription of mature mRNA and subsequent integration at random sites of a genome (Hurles 2004). The resulting duplicated genes tend to lack introns and have poly(A) tracts at their 3’ ends. All sequenced MgMEP genes have the same introns and lack poly(A) tracts, indicating that retrotransposition is an unlikely mechanism for generating these paralogs. The other possible mechanism includes unequal crossing over resulting from homologous recombination between DNA repeat sequences. Recent evidence has shown an abundance of Alu elements (a dispersed repetitive ~300 bp long, constituting ~10% of human genome) at the junctions between duplicated and single copy sequences, indicating that the duplications are likely to have resulted from unequal crossing over between these repeats (Bailey et al. 2003). The MgMEP genes may have been duplicated by unequal crossing over between retrotransposons flanking the genes, followed by translocation of the duplicate sequence into another genomic location.
An unusually high frequency of C:G to T:A transitions among MGLR-4 and MGLR-5 elements can result from Repeat-Induced Point mutation (RIP) or RIP-like processes that were originally identified in *Neurospora crassa* (CAMBARERI et al. 1989; GALAGAN and SELKER 2004; SELKER et al. 1987). In fungi, these processes have been proposed as one of the defense mechanisms against DNA repeat sequences including transposable elements. The C:G to T:A transitions are likely to occur through deamination of 5-methyl cytosine to generate thymine, which subsequently leads to change guanine in the complementary strand to adenine. The main features of the original RIP mechanism found in *N. crassa* include: (i) generating mutations that are exclusive C:G to T:A transitions into both copies of duplicated sequences; (ii) occurring preferentially in CpA dinucleotides; (iii) being triggered by duplicated sequences (>400 bp with >80% sequence identity) generated by chromosomal duplications or by transformation experiments; and (iv) acting during the sexual cycle. Since its first discovery in *N. crassa*, RIP and RIP-like processes have been reported in *Podospora anserina* (GRAIA et al. 2001), *Aspergillus fumigatus* (NEUVEGLISE et al. 1996), *Aspergillus nidulans* (NIELSEN et al. 2001), *Fusarium oxysporum* (HUA-VAN et al. 2001), *M. grisea* (IKEDA et al. 2002; NAKAYASHIKI et al. 1999; THON et al. 2004), and *Leptosphaeria maculans* (IDNURM and HOWLETT 2003). The sexual stage of *M. grisea* has only been demonstrated under laboratory conditions, but never in nature (KATO et al. 2000). The presence of a RIP-like process suggests that sexual recombination has occurred at some time in the past.
A DUAL SELECTION BASED, TARGETED GENE REPLACEMENT TOOL FOR MAGNAPORTHE GRISEA AND FUSARIUM OXYSPORUM

Chang Hyun Khang1, Sook-Young Park1, Yong-Hwan Lee2, and Seogchan Kang1
1Department of Plant Pathology, The Pennsylvania State University, University Park, PA 16802, USA; 2School of Agricultural Biotechnology, Seoul National University, Seoul 151-742, Korea

Note: This chapter has been submitted for publication.

4.1 Introduction

Better understanding of fungal biology has greatly increased in recent years as a result of significant advances in fungal genomics, including the release of full fungal genome sequences and the development of efficient molecular tools. Among these tools, transformation-mediated mutagenesis has undoubtedly been the most widely applied method for studying gene function in fungi. In most filamentous fungi, transformation results from the integration of the transforming DNA into the fungal genome by either non-homologous or homologous recombination. Transformation with a plasmid through a non-homologous recombination pathway has been widely used for random mutagenesis in fungi, as an alternative to chemical or radiation mutagenesis. The integration of a plasmid into a gene provides a convenient molecular tag to rescue and characterize this gene (MULLINS and KANG 2001). Transformation through a homologous recombination pathway permits targeted gene replacement and requires a plasmid carrying two DNA fragments from the targeted locus of sufficient length to promote homologous recombination (HAMER et al. 2001). An important advance in fungal transformation has been the recent development of Agrobacterium tumefaciens-mediated transformation

Fungi with a low frequency of homologous recombination require that a large number of transformants be generated and screened to identify the desired mutant (MULLINS and KANG 2001). To circumvent this time-consuming process, we developed a targeted gene replacement method (termed ATMT-DS) based on ATMT followed by a dual (positive and negative) selection (DS) of transformants. We have evaluated the factors affecting the efficiency of targeted gene replacement using ATMT-DS in two plant pathogenic fungi, *Magnaporthe grisea* and *Fusarium oxysporum*. *M. grisea*, the causal agent of rice blast disease, is the most damaging pathogen of rice. In addition to its economic significance, *M. grisea* is one of the main models for understanding fungal pathogenicity (KANG et al. 2000; TALBOT et al. 1993). The *F. oxysporum* complex comprises diverse, largely soilborne fungi including plant pathogens causing vascular wilt on more than 100 cultivated plant species. We showed that ATMT-DS is an efficient tool for targeted gene replacement in *M. grisea* and *F. oxysporum*, and is likely to be applicable to manipulating genes in other filamentous fungi.
4.2 Materials and Methods

4.2.1. Strains, media, and ATMT-DS

*Agrobacterium tumefaciens* strains AGL1 and EHA105 (Klee 2000) were used to transform *M. grisea* strains KJ201 (Park et al. 2000), and 4091-5-8 (Valent et al. 1986), and *F. oxysporum* O-685 (Mullins et al. 2001). The following fungal and oomycete strains tested for sensitivity to 5-fluoro-2’-deoxyuridine (F2dU) were from our laboratory (*Verticillium dahliae* and *Crinipellis perniciosa*) or our colleagues at Penn State University (University Park, PA), including David Geiser (*Aspergillus oryzae, A. fumigatus, A. nidulans*), Hye-Ji Kim (*Thielaviopsis* sp.), Wakar Uddin (*Rhizoctonia solani*), and Gary Moorman (*Botrytis cinerea, Pythium aphanidernatum, Pythium ultimum, Pythium irregulare, Phytophthora cactorum, and Phytophthora cinnamomii*). Ganciclovir and F2dU (Sigma), hygromycin B (Calbiochem), and geneticin (Sigma) were dissolved in water and filter-sterilized to prepare stock solutions and stored at –20°C except hygromycin B (4°C). For testing sensitivity to ganciclovir and F2dU, *M. grisea* was grown on complete medium (Valent et al. 1986). Potato dextrose agar (Difco) was used for testing other fungi and oomycetes. ATMT was performed as previously described (Mullins and Kang 2001). Different amounts of hygromycin B were used for selecting transformants (250 µg/ml for *M. grisea* and 50 µg/ml for *F. oxysporum*). Direct selection of putative gene replacement mutants was carried out by regenerating transformants in the presence of both F2dU (5 µM or 50 µM) and hygromycin B. Different amounts of geneticin (800 µg/ml for *M. grisea* and 50 µg/ml for *F. oxysporum*) were utilized to identify transformants expressing *neo* (a gene conferring resistance to geneticin). For determining the presence of *HSVtk* (a herpes simplex virus thymidine kinase gene) and mutant allele in transformants, fungal genomic DNA was analyzed by PCR and/or Southern hybridization.
4.2.2 Vector construction

The ChGPD-HSVik construct (1.8kb EcoRI-HindIII fragment) in pGEM3Zf (Promega) consists of three modules: the Cochliobolus heterostrophus glyceraldehyde-3-phosphate dehydrogenase (ChGPD) promoter (0.5kb EcoRI-BamHI fragment), the open reading frame (ORF) of HSVik (1.1kb BamHI-Sall fragment), and the Neurospora crassa β-tubulin gene terminator (0.2kb SphI-HindIII fragment). The HSVik ORF was amplified by PCR from pPNT (TZIFIRA et al. 2004) using a pair of primers containing appropriate restriction sites at their 5’ end. The promoter and terminator modules were amplified using pSM565 (BOURETT et al. 2002) as a template. All the modules were sequenced to verify their sequence.

Plasmid pBHt2-tk was constructed by cloning the 1.8kb EcoRI-HindIII fragment carrying ChGPD-HSVik between the EcoRI and HindIII sites of pBHt2 (MULLINS et al. 2001). To construct pGKO1, the 1.8kb EcoRI-HindIII fragment was made blunt by treating it with Klenow fragment in the presence of dNTPs, and cloned between the blunted XhoI and BstXI sites of pCAMBIA1300 (http://www.cambia.org.au).

Two genes, F. oxysporum FoSNF1 (OSPINA-GIRALDO et al. 2003) and M. grisea MHP1, a gene encoding a class II hydropphin (Y. Lee, unpublished data; M. grisea genome contig 2.206 and locus ID MG01173.3), were utilized to test factors affecting the efficiency of ATMT-DS. To produce pGKO1-fosnf1, a 1kb fragment corresponding to FoSNF1 (covering 322bp of the promoter region and 683bp of the open reading frame) was amplified from F. oxysporum O-685 by PCR using the following primers (Fig. 6): SNF-f (5’-AGCACTAGTAAATCTACCCGAGGCCAGTC-3’) and SNF-r (5’-AGGCAATTGGGCGATTTTGACGTGAGA-3’). The underlined sequences in SNF-f and SNF-r correspond to SpeI and MfeI sites, respectively. After cloning the amplified fragment into pGEM-T Easy (Promega), a 60bp HindIII-HindIII fragment of the amplified FoSNF1 (Fig. 6) was replaced with the 1.4kb HpaI fragment of pBC1004 carrying hpt, a gene encoding hygromycin B phosphotransferase (CARROLL et al. 1994). The resulting mutant allele was digested with SpeI and MfeI and cloned between the EcoRI and XbaI sites of pGKO1 to produce pGKO1-fosnf1.
For pGKO1-\textit{mhp1}, a 1.5kb fragment containing \textit{MHP1} (starting at 558bp upstream of the start codon and ending at 983bp downstream from the start codon) was amplified from \textit{M. grisea} 70-15 by PCR using the following primers (Fig. 6): MHP-f (5’-ACGGAATTTCTCGACATGGACCGTCTTG-3’) and MHP-r (5’-AGCTCTAGAGTACCAAGCCGCACCAC-3’). The underlined sequences in MHP-f and MHP-r correspond to EcoRI and XbaI sites, respectively. The \textit{hpt} gene was inserted into the blunted BglII site (212bp downstream from the start codon) located in the middle of the amplified \textit{MHP1} locus to generate an \textit{mhp1} mutant allele. The resulting mutant allele was digested with EcoRI and XbaI and cloned between the EcoRI and XbaI sites of pGKO1 to produce pGKO1-\textit{mhp1}. The \textit{mhp1} mutant allele in pGKO1-\textit{mhp1} was isolated as an EcoRI-HindIII fragment and cloned in pGKO2 to produce pGKO2-\textit{mhp1}.

\textbf{Figure 6.} Schematic diagram of the \textit{Fusarium oxysporum} \textit{FoSNF1} and \textit{Magnaporthe grisea} \textit{MHP1} loci. Exons and introns of the \textit{FoSNF1} and \textit{MHP1} genes are represented by open boxes and hooks, respectively. Numbers indicate the genomic positions of relevant restriction sites/markers relative to the translation start codon (A of ATG as +1). PCR primers that were used to amplify each locus for mutant allele construction were denoted by arrow. The filled triangle indicates the insertion site of the \textit{hpt} gene to create a mutant allele for each gene (between two HindIII sites for the \textit{FoSNF1} mutant allele and at the BglII site for the \textit{MHP1} mutant allele).
To generate two binary vectors, pDHt-KS and pDHt-SK (identical except the orientation of their MCS), a 300bp PvuII-PvuII fragment of pDHt (MULLINS and KANG 2001), containing the multiple cloning site (MCS) was replaced with a 0.8kb HpaI-StuI fragment isolated from pGreenII0000 (HELLENS et al. 2000).

For constructing pNHTK and pTKHN, three selectable markers, neo (1.2kb BamHI-SalI fragment), hpt (1.4kb SalI-EcoRI fragment), and ChGPD-HSVtk (1.8kb EcoRI-HindIII fragment), were initially cloned between BamHI and HindIII sites of pBluescript SK (Stratagene) in the order of SpeI-BamHI-neo-hpt-ChGPD-HSVtk-HindIII, resulting in pSK1697. The 4.4kb SpeI-HindIII fragment of pSK1697 was cloned between the SpeI and HindIII sites of pDHt-SK and pDHt-KS to generate pNHTK and pTKHN, respectively.

Selected restriction sites on the ChGPD-HSVtk construct in pGEM3Zf were mutagenized using QuikChange Multi Site-Directed Mutagenesis kit (Stratagene) according to the manufacturer’s instruction. The modified ChGPD-HSVtk construct (as a blunted EcoRI-HindIII fragment) was cloned into a blunted SacI site of pDHt-KS, resulting in pGKO2. To allow for cloning of mutant allele into pGKO2 without relying on available restriction sites, we constructed pGKO2-Gateway as follows: the ccdB (control of cell death B) and chloramphenicol resistance genes flanked by the λ, attP sites in pDONR201 (Invitrogen) was amplified by PCR using the following primers: 5’-TCGCTCTAGAAATAATGATTTTATTTGAC-3’ and 5’-TCGCAAGCTTGGATGGCAAATAATGAT-3’ (the underlined sequences correspond to XbaI and HindIII sites, respectively). The resulting product (2.3kb) was first cloned in pGEM-T Easy for sequence verification and was subsequently cloned between the XbaI and HindIII sites of pGKO2.
4.3 Results

4.3.1 Herpes simplex virus thymidine kinase (HSVtk) functions as a negative selection marker in diverse fungi

ATMT-DS is designed to counter-select ectopic transformants using a gene conferring lethality when expressed in transformants (the negative selection marker). This marker flanks a mutant allele that contains a positive selection maker such as the hygromycin B resistance gene. Both markers are located between LB and RB of the T-DNA. Ectopic transformants express both the negative and positive selection markers, while transformants resulting from gene replacement lack the negative selection marker. Therefore, the negative selection should eliminate ectopic transformants (Fig. 7), facilitating the identification of transformants resulting from gene replacement since they are resistant to this negative selection. We tested two genes, Dtx-A encoding diphtheria toxin subunit A, and HSVtk encoding a viral thymidine kinase, as potential negative selection markers for fungi. Although Dtx-A has been successfully utilized as a negative selection marker in plants (Czako and An, 1991; Terada et al., 2002) and mammalian cells (Yagi et al., 1990), Dtx-A expressed using two different fungal promoters did not appear to be toxic to M. grisea and F. oxysporum (data not shown). The HSVtk gene product converts nucleoside analogs, such as ganciclovir and F2dU, to toxic compounds and has been utilized as a conditional negative selection marker in diverse organisms (Capecchi, 1989; Chen et al., 2002; Duraisingham et al., 2002; Sachs et al., 1997).
Figure 7. Schematic diagram of ATMT-DS. *Agrobacterium tumefaciens* cells, carrying a binary vector that contains a mutant allele (disrupted by a positive selection marker, such as *hpt* conferring resistance to hygromycin B; marked as the filled box) and *HSVtk* (encoding thymidine kinase that convert nucleoside analogs such as 5-fluoro-2'-deoxyuridine (F2dU) to a compound toxic to fungi; denoted by the diamond) on the T-DNA, are co-incubated with fungal cells in the presence of acetosyringone (AS), a chemical inducer of virulence genes of *A. tumefaciens*. During co-cultivation, DNA situated between the left border (LB) and right border (RB) of the T-DNA is transported into fungal nuclei (probably as a complex with *A. tumefaciens* Vir proteins denoted by the circles). Homologous recombination between the native gene and the mutant allele on the T-DNA leads to the loss of *HSVtk*. If the T-DNA integrates into a random location in the fungal genome via non-homologous recombination, both *hpt* and *HSVtk* will be expressed. Targeted gene replacement mutants can be selected by subjecting transformants to both the positive (hygromycin B) and negative (F2dU) selection agents.

Figure 8. Schematic diagrams of the T-DNA regions in the binary vectors constructed in this study. The LB and RB of the T-DNA are denoted by vertical lines. The orientation of transcription from *hpt* and *HSVtk* is indicated by arrow (5’ to 3’). MCS1 corresponds to the multiple cloning site of pCAMBIA1300. The multiple cloning site of pGreenII0000 cloned in pDHt was designated as MCS-SK or MCS-KS depending on its orientation: *KpnI* (K) and *SacI* (Sc) sites are shown to indicate the orientation of the MCS1 relative to other markers. A modified version of *HSVtk* obtained by site-directed mutagenesis is denoted as *HSVtk(M)*. Gateway corresponds to the *ccdB* and chloramphenicol-resistance genes flanked by the *λattP* sites. Drawings are not to scale.
Transformants of *Aspergillus fumigatus*, *Botrytis cinerea*, *M. grisea* and *F. oxysporum* generated using pBHt2-tk (Fig. 8), a binary vector carrying the HSVtk gene under the control of the *Cochliobolus heterostrophus* glyceraldehyde-3-phosphate dehydrogenase (*ChGPD*) gene promoter and the *Neurospora crassa* β-tubulin gene terminator on the T-DNA, exhibited sensitivity to ganciclovir (with the exception of *B. cinerea*) and F2dU. Sensitivity to F2dU was much greater than that to ganciclovir (Fig. 9 & not shown). For instance, the effective concentration of F2dU for completely inhibiting the growth of *M. grisea* was approximately 0.5 µM, while 1 mM ganciclovir was needed to achieve the same growth inhibition. Ganciclovir failed to inhibit the growth of *B. cinerea* transformants even at 2 mM, while 5 nM of F2dU was sufficient to inhibit their growth (Fig. 9D). HSVtk transformants of *A. fumigatus* were much less sensitive to F2dU than were *B. cinerea*, *M. grisea* and *F. oxysporum* transformants, requiring 50 µM F2dU for significant growth inhibition.

**Figure 9.** Growth of *Magnaporthe grisea*, *Fusarium oxysporum*, *Aspergillus fumigatus*, and *Botrytis cinerea* in the presence of 5-fluoro-2’-deoxyuridine (F2dU) or ganciclovir. Wild-type strains (wt) and transformants with *ChGPD-HSVtk (tk)* of *M. grisea* (*A & E*), *F. oxysporum* (*B*), *A. fumigatus* (*C*), and *B. cinerea* (*D*) were grown in the presence of F2dU (*A-D*) or ganciclovir (*E*) at concentrations ranging from 5 nM to 50 µM (F2dU) or 1 µM to 2 mM (ganciclovir).
In contrast, wild-type strains of these and other fungal and oomycete species, including ascomycetes (*Aspergillus oryzae*, *A. nidulans*, *Thielaviopsis* spp., and *Verticillium dahliae*), basidiomycetes (*Rhizoctonia solani* and *Crinipellis perniciosa*), and oomycetes (*Pythium aphanidermatum*, *P. ultimum*, *P. irregulare*, *Phytophthora cactorum*, and *P. cinnamomi*), did not exhibit sensitivity to F2dU or ganciclovir at the concentrations that completely blocked the growth of *HSVtk* transformants (data not shown), suggesting the applicability of *HSVtk* as a negative selection marker for ATMT-DS in diverse fungal and oomycete species.

4.3.2 Targeted gene replacement in *F. oxysporum* and *M. grisea* genes using ATMT-DS

We utilized two genes, *F. oxysporum* *FoSNF1*, a gene playing a critical role in carbon catabolite repression and pathogenicity (Ospina-Giraldo et al. 2003) and *M. grisea* *MHP1*, a gene encoding class II hydrophobin (Y. Lee, unpublished data), to test ATMT-DS (Fig. 6). To determine the degree of enrichment of gene replacement mutants via ATMT-DS, we first isolated hygromycin B-resistant transformants in the absence of F2dU and subsequently analyzed for their sensitivity to F2dU and the presence of target mutation (Table 4). In addition, to determine if bacterial and/or fungal strain-specific factors affect the efficiency of targeted gene replacement, we introduced gene disruption vectors pGKO1-*fosnf1* and pGKO1-*mhp1* into two different *A. tumefaciens* strains, AGL1 and EHA105, and also employed two strains of *M. grisea*, KJ201 (Park et al. 2000) and 4091-5-8 (Valent et al. 1986). With *F. oxysporum*, AGL1 yielded a higher gene replacement frequency than did EHA105 (20% vs. 9%). With *M. grisea*, AGL1 was better than EHA105 in generating gene replacement mutants in 4091-5-8 (6% vs. 0%) but produced fewer gene replacement mutants in KJ201 than did EHA105 (26% vs. 51%). With both AGL1 and EHA105, the frequencies of gene replacement in KJ201 were significantly higher than those in 4091-5-8. In both fungal species, some hygromycin B and F2dU-resistant transformants had intact *MHP1* or *FoSNF1* genes and likely resulted from ectopic integrations of the T-DNA that failed to express *HSVtk*. The frequency of
such false positive (FP) ranged from 9-50% in *F. oxysporum* to 24-100% in *M. grisea* (Table 4). In *M. grisea*, KJ201 yielded lower frequencies of FP than did 4091-5-8.

**Table 4. Efficiency of targeted gene replacement via ATMT-DS in *M. grisea* and *F. oxysporum***

<table>
<thead>
<tr>
<th>Clones used</th>
<th>Fungal strain†</th>
<th>A. tumefaciens strain</th>
<th>HR²</th>
<th>HFR³</th>
<th>TGR⁴</th>
<th>False positive (FP)⁵</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGKO1-fosnf1</td>
<td>O-685</td>
<td>AGL1</td>
<td>51</td>
<td>11</td>
<td>10 (20%)</td>
<td>1 (9%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EHA105</td>
<td>34</td>
<td>6</td>
<td>3 (9%)</td>
<td>3 (50%)</td>
</tr>
<tr>
<td>pGKO1-mhp1</td>
<td>KJ201</td>
<td>AGL1</td>
<td>70</td>
<td>26</td>
<td>18 (26%)</td>
<td>8 (31%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EHA105</td>
<td>49</td>
<td>33</td>
<td>25 (51%)</td>
<td>8 (24%)</td>
</tr>
<tr>
<td></td>
<td>4091-5-8</td>
<td>AGL1</td>
<td>31</td>
<td>10</td>
<td>2 (6%)</td>
<td>8 (80%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EHA105</td>
<td>40</td>
<td>8</td>
<td>0 (0%)</td>
<td>8 (100%)</td>
</tr>
</tbody>
</table>

† *Fusarium oxysporum* strain O-685 was transformed using pGKO1-fosnf1, and *M. grisea* strains KJ201 and 4091-5-8 were transformed using pGKO1-mhp1.

² Total number of hygromycin B-resistant (HR) transformants isolated from two to four independent transformations.

³ The number of HR transformants resistant to 5 µM F2dU (HFR).

⁴ The number and percentage of targeted gene replacement (TGR) mutants among HR.

⁵ The number and percentage of false positive (FP) transformants (resistant to both hygromycin B and F2dU but do not carry the targeted gene replacement) among HFR.

To determine whether FP transformants resulted from the truncation of HSVtk during ectopic integration of the T-DNA, a set of 28 FP transformants, including three from *F. oxysporum* and 25 from *M. grisea*, were analyzed by PCR and Southern hybridization, (Fig. 10). All FP transformants from *F. oxysporum* (3/3) and KJ201 (13/13), and most FP transformants from 4091-5-8 (9/12) corresponded to ectopic transformants with various truncations of LB and HSVtk (Fig. 10). One 4091-5-8 FP had extensive truncations at both the LB and RB. The remaining two FP transformants from 4091-5-8, however, had intact LB and ChGPD-HSVtk but were insensitive even to 50 µM F2dU (data not shown), suggesting that the expression of HSVtk was suppressed probably due to the chromosomal context of inserted T-DNA.
Figure 10. Southern analysis of *M. grisea* ATMT-DS transformants. The hatched box interrupted by *hpt* denotes the *mhp1* mutant allele cloned in pGKO1. The diagram of the T-DNA was not drawn to scale. *Eco*RI-digested genomic DNA of wild type 4091-5-8 strain (lane 1) and its transformants, including one gene replacement mutant (lane 2), one ectopic transformant (lane 3), and three different types of false positive transformants (lanes 4-6), was hybridized with each of the four probes shown underneath the T-DNA diagram: A, 0.3kb fragment covering the region between the LB and the *ChGPD* promoter; B, 0.4kb fragment covering the *ChGPD* promoter; C, 2.9kb fragment covering both *hpt* and parts of the *MHP1* locus; D, 250bp fragment covering the region between the RB and the mutant allele. The arrow in Fig. 10C marks the wild-type *MHP1* gene, which is absent in the gene replacement mutant (lane 2).

When putative gene replacement mutants were directly selected by regenerating transformants of O-685, KJ201, and 4091-5-8 in the presence of both hygromycin B and 5 µM F2dU, in all strains, the negative selection was leaky: a significant fraction of transformants (71%, 20%, and 82% in O-685, KJ201, and 4091-5-8, respectively) exhibited sensitivity to F2dU when transferred to fresh media containing the same concentration of F2dU (data not shown). During selection, *A. tumefaciens* cells began to lyse, possibly releasing nucleosides counteracting the F2dU toxicity. If so, increasing the concentration of F2dU should reduce the leakiness of negative selection. However, even 50 µM F2dU did not alleviate the problem, as a significant number of O-685 and 4091-5-8 transformants still remained sensitive to 5 µM F2dU (Table 5). During this experiment, we also noticed that in both fungi the presence of 50 µM F2dU consistently reduced (2 to
4 fold) the number of transformants compared to number obtained with 0 or 5 µM F2dU, suggesting that high F2dU interfere with the efficiency of transformation.

### Table 5. Leakiness of direct negative selection during the regeneration of transformants

<table>
<thead>
<tr>
<th>Fungal strain</th>
<th>Direct selection</th>
<th>F2dU-insensitive transformants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F2dU (µM)</td>
<td>Number of transformants</td>
</tr>
<tr>
<td>O-685</td>
<td>0</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>83</td>
</tr>
<tr>
<td>4091-5-8</td>
<td>0</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>41</td>
</tr>
</tbody>
</table>

1 *Fusarium oxysporum* strain O-685 and *M. grisea* strain 4091-5-8 were transformed using pGKO1-*fosnf1* and pGKO1-*mhp1*, respectively.

2 Total number of transformants isolated from selection plates containing both hygromycin B and F2dU (0 µM, 5 µM or 50 µM). Seven plates were used for each treatment.

3 The number and percentage of primary transformants that grew on a new plate containing 5 µM F2dU.

4.3.3 Stability of *HSVtk* depends on fungal strains and its location in the T-DNA

In pGKO1-*fosnf1* and pGKO1-*mhp1* gene replacement vectors, *HSVtk* was located at LB of the T-DNA. We have observed a significant number of transformants with truncations of LB including *HSVtk* (Fig. 10). To evaluate the stability of markers according to their location in the T-DNA, we constructed two vectors carrying neo and *HSVtk* on both sides of the *hpt* positive selection marker (Fig. 11). In the pTKHN vector, *HSVtk* is located near LB and neo near RB. In the pNHTK vector, neo is located near LB and *HSVtk* near RB. Hygromycin B-resistant transformants of *F. oxysporum* and *M. grisea* obtained with either pNHTK or pTKHN (Fig. 11) were scored for their resistance to F2dU (loss of *HSVtk*) and geneticin (presence of neo). *HSVtk* was more easily lost or inactive when located at LB (4-13%) than at RB (1-2%) in the three fungal strains tested (Fig. 11). There was no apparent positional difference for neo loss in *M. grisea* while in *F. oxysporum* the stability of neo followed the same trend as *HSVtk*. In *F. oxysporum*, the loss of neo was two times higher when the marker was located at LB compared to RB.
The fact that neo was significantly more prone to loss of function than HSVtk at both LB and RB locations in M. grisea, is puzzling.

![Diagram](image)

**Figure 11.** Stability of neo and HSVtk located at either LB or RB of the T-DNA. Each fungal strain was transformed using (A) pTKHN and (B) pNHTK. The total number of hygromycin B-resistant (HR) transformants analyzed, and the number and percentage of HR sensitive to geneticin and F2dU (Loss of neo), resistant to F2dU and geneticin (Loss of HSVtk), and sensitive to geneticin and resistant to F2dU (loss of both markers) were indicated in the tables.

### 4.3.4 Construction of improved targeted gene replacement vectors

To facilitate ATMT-DS mediated targeted gene replacement, we constructed a number of improved vectors (Fig. 8). The ChGPD-HSVtk construct on pGKO1 contains one or more of the following restriction sites: BamHI, EcoRV, PstI, SacI, SalI, and SmaI. Most of these sites (except EcoRV) are also present in the multiple cloning site of
pGKO1, significantly reducing the number of available cloning sites. We removed these restriction sites from the \textit{ChGPD-HSVtk} construct via site-directed mutagenesis, and at the same time, improved codons at the mutated sites based on the fungal codon usage (http://www.kazusa.or.jp/codon). To further expand the number of available restriction sites for cloning, we replaced the original multiple cloning site of the binary vector pDHt (\textit{Mullins et al. 2001}) with the multiple cloning site from pGreenII0000 (\textit{Hellers et al. 2000}), resulting in pDHt-KS and pDHt-SK with 15 unique restriction sites.

The stability of \textit{HSVtk} was significantly higher at RB than that at LB (Fig. 11). To reduce the generation of FP transformants caused by loss of the \textit{HSVtk} gene during T-DNA integration, the mutated \textit{ChGPD-HSVtk} construct was cloned at the SacI site (located near RB) in the multiple cloning site of pDHt-KS to generate pGKO2 (Fig. 8). The frequency of FP with pGKO2 was compared with that with pGKO1 using \textit{M. grisea}. We cloned \textit{mhp1} in pGKO2 to produce pGKO2-\textit{mhp1}. After transforming \textit{M. grisea} strain KJ201 with pGKO1-\textit{mhp1} and pGKO2-\textit{mhp1}, we counted the number of FP transformants among randomly picked transformants (132 transformants for each construct). With pGKO1-\textit{mhp1}, 56% of the transformants resistant to both hygromycin B and F2dU (14/25) were FP, whereas with pGKO2-\textit{mhp1}, the frequency of FP was only 6% (2/35). A derivative of pGKO2, termed pGKO2-Gateway, was also produced by introducing the GATEWAY™ system (Stratagene), designed to facilitate the movement of DNA fragments between vectors through the use of lambda recombinase instead of restriction enzymes and ligase, into pGKO2 (Fig. 8).

4.4 Discussion

A significant number of fungal genomes have been sequenced or are currently being sequenced (\textit{Galagan and Selker 2004; Goffeau et al. 1996}) (http://www.genomesonline.org/). In many fungi, determining gene function through the generation of null mutants by targeted gene replacement is limited by the low efficiency of homologous recombination. Therefore, the development of a technique to circumvent this limiting factor is critical for fungal functional genomics. ATMT has many
advantages for gene manipulation in fungi, including high transformation efficiency (De Groot et al. 1998; Mullins and Kang 2001; Rho et al. 2001), increased frequency of homologous recombination (Bundock et al. 1999; Dobinson et al. 2004; Zwiers and De Waard 2001), and ability to transform intact spores and hyphae (De Groot et al. 1998). To further improve targeted gene replacement based on ATMT, we constructed and tested a binary Agrobacterium vector (pGKO1) carrying as a negative selection marker against ectopic transformants the HSVtk gene. This gene was originally used in animal cells to facilitate targeted gene replacement (Capecchi 1989). In Neurospora crassa, a different negative selection marker, the N. crassa mat a-1 gene conferring heterokaryon incompatibility in the strains carrying the mat A gene, was utilized to enhance the efficiency of gene replacement through protoplast-mediated transformation (Pratt and Aramayo 2002). While the N. crassa mat a-1 gene allowed a significant enrichment of gene replacement mutants, its utility is limited because this marker confers toxicity only to certain strains of N. crassa. In contrast, HSVtk appears to function as a universal, conditional negative selection marker. In addition to the four fungal species tested in our study (Fig. 9), transformants of N. crassa (Pratt and Aramayo 2002; Sachs et al. 1979), Leptosphaeria maculans (Gardiner and Howlett 2004), and the human pathogenic basidiomycete Cryptococcus neoformans (Y. Chang and J. Kwon-Chung at NIH, personal communication) that express HSVtk also exhibited sensitivity to F2dU. Our survey suggested the lack of an enzyme equivalent to HSVtk in many other fungi and oomycetes, further supporting the utility of HSVtk as a universal negative selection marker. Only a wild-type strain of B. cinerea exhibited noticeable sensitivity to F2dU (Fig. 9).

A targeted gene replacement strategy identical to ATMT-DS was recently applied to disrupt genes in L. maculans (Gardiner and Howlett 2004). Our work further extended this study as follows: (i) We demonstrated that HSVtk could function as a conditional negative selection marker in diverse fungal/oomycete species. Given that diverse fungi have now been successfully transformed via ATMT, ATMT-DS can be broadly adopted to disrupt fungal genes; the only modification that might be needed would be to replace the ChGPD promoter driving the expression of HSVtk with an appropriate promoter for
certain target fungi. Due to the modular structure of the \textit{HSVtk} construct, such a modification should be simple. (ii) We showed that the stability of \textit{HSVtk} varied according to its location in the T-DNA and the fungal species/strains transformed (see below for more discussion). (iii) By employing two different \textit{A. tumefaciens} strains and two strains of \textit{M. grisea}, we tested if bacterial and/or fungal strain-specific factors affect the efficiency of targeted gene replacement through ATMT-DS (Table 4).

Although ATMT-DS facilitated the rapid identification of gene replacement mutants in \textit{F. oxysporum} and \textit{M. grisea} through the efficient counter-selection of ectopic transformants, we encountered two limiting factors for ATMT-DS. First, we identified FP transformants in both \textit{F. oxysporum} and \textit{M. grisea} that are resistant to hygromycin B and F2dU but do not result from targeted gene replacement (Table 4). These FP transformants mainly resulted from ectopic integrations of the T-DNA truncated at their LB leading to the loss of \textit{HSVtk} (Fig. 10). Although RB was not immune to truncation, when \textit{HSVtk} was located near RB, the frequency of its loss (or inactivation) was significantly lower (ranging from <1\% in \textit{F. oxysporum} to 2\% in \textit{M. grisea} 4091-5-8) than that near LB (ranging from 4\% in \textit{F. oxysporum} to 13\% in \textit{M. grisea} 4091-5-8) (Fig. 11). In plants, when truncations in inserted T-DNAs occur, they appear more extensive at LB than RB (Tzfira et al. 2004).

Differential stability of the T-DNA borders is likely caused by different molecular changes occurring at LB and RB during T-DNA integration (Tzfira et al. 2004). In plants, T-DNA integration seems to be initiated by micro-homology dependent annealing of the LB sequence or an adjacent region to a region in the host genome. The T-DNA sequence at the 3‘ side of the annealed region is trimmed by exonucleases prior to the ligation of the T-DNA to the host DNA, suggesting that initial annealing through a region in the middle of the T-DNA can bring about extensive T-DNA deletion. The RB side of T-DNA single-strand, which is covalently attached to the \textit{A. tumefaciens} VirD2 protein, may also anneal to the host genome but be protected by VirD2 until this end is ligated (presumably mediated by VirD2 and/or unknown host ligases attracted by VirD2) to the host DNA. Considering that ATMT of fungi seems to involve host and bacterial factors that are also important for successful ATMT of plants (Tzfira et al. 2004), integration
and truncation of the T-DNA in fungi may also be caused by the same mechanism observed in plants.

To reduce the number of FP transformants, we constructed a new binary vector, pGKO2 (Fig. 8). Targeted gene replacement using pGKO2 indeed significantly reduced the frequency of FP in *M. grisea* (56% FP with pGKO1-<i>mhp1</i> vs. 6% FP with pGKO2-<i>mhp1</i>). Vector pGKO2-Gateway was also constructed to facilitate the cloning of fungal genes for targeted mutagenesis. Although this improvement should alleviate the problem of FP in *M. grisea* and *F. oxysporum*, in fungal species that exhibit a high rate of T-DNA truncations and a low homologous recombination frequency, FP transformants still remain a significant problem. This problem seems particularly severe in *L. maculans*, since the use of a binary vector carrying *HSVtk* located at both LB and RB, still led to a significant number of FP transformants (GARDINER and HOWLETT 2004). In the long run, solving this problem will require a better understanding of the mechanism underpinning ATMT and the gene replacement using the T-DNA. Further studies on such a mechanism will not only help us improve ATMT-DS to allow efficient gene replacements even in fungi like *L. maculans* but also complement efforts to understand the mechanism of ATMT in plants (TZFIKA et al. 2004). Studies for determining the bacterial and fungal factors that affect ATMT have been conducted using *Saccharomyces cerevisiae* (BUNDOCK et al. 1995; VAN ATTIKUM and HOONYAAS 2003) and *Aspergillus awamori* (MICHELSE et al. 2004a).

Another problem, although minor, is the leakiness of the negative selection during the regeneration of transformants. Considering that even 50 µM F2dU failed to effectively select against F2dU-sensitive transformants during regeneration, it seems unlikely that nucleosides released from dead *A. tumefaciens* cells are responsible for the leakiness. Alternatively, we hypothesize that the expression of *HSVtk* driven by the *ChGPD* promoter might be suppressed during regeneration. If so, using a different fungal promoter might solve the problem. For now, screening transformants for their sensitivity to F2dU after their regeneration is sufficient for solving this problem.

A number of factors seem to affect the efficiency of isolating gene replacement mutants through ATMT-DS (Table 4 and Fig. 11). As discussed above, the position of
HSVtk on the T-DNA is one such factor. Interestingly, the stability of HSVtk located at LB was significantly higher in *F. oxysporum* (2-3 folds) than in *M. grisea* (Fig. 11). This result is consistent with those obtained with pGKO1-*fosnf1* and pGKO1-*mhp1* gene replacement vectors (Table 4), as we observed significantly less FP transformants in *F. oxysporum* than in *M. grisea*. Therefore, the truncation of LB during T-DNA integration appeared more frequent in *M. grisea* than in *F. oxysporum*. In *M. grisea*, the stability of HSVtk located at LB was significantly higher in KJ201 than 4091-5-8 (Table 4 and Fig. 11). These results suggest that a significant role of the host genotype in the control of this aberrant integration process. Regardless of *A. tumefaciens* strains used, the frequency of gene replacement in KJ201 was significantly higher than that in 4091-5-8. Slight sequence differences between the mutant allele and its chromosomal copy in 4091-5-8 might have caused the reduced gene replacement frequency; the *mhp1* allele used for mutagenesis is identical to that of KJ201, but exhibits a number of polymorphic sites (31 out of 1540bp) to that of 4091-5-8 (data not shown). Considering that a number of factors can potentially affect the efficiency of targeted gene replacement via ATMT-DS, for a new fungal species to be mutagenized, prior to launching a large-scale mutagenesis experiment, carefully evaluating such factors for their role in the efficiency of gene replacement is recommended.

### 4.5 Acknowledgements

We thank our colleagues for sharing their strains, vectors, and/or unpublished data. We acknowledge Kathy Dobinson, Gretchen Kuldau, and Justin Dillon for critically reviewing the manuscript. We also greatly appreciate many helpful suggestions by the anonymous reviewers of the manuscript. This work was supported in part by grants from USDA-NRI (2002-02367; SK), Crop Functional Genomics Center of the 21st Century Frontier Research Program (CG1413; YHL), and the Korea Science and Engineering Foundation (Agricultural Plant Stress Research Center; YHL).
Chapter 5

GENERAL DISCUSSION

Magnaporthe grisea causes rice blast disease, one of the most devastating crop diseases worldwide. Effective control of the blast disease using host resistance has not been effective, due to the frequent appearance of new races with ability to infect previously resistant cultivars. Besides its economic importance, rice blast has become one of the major model systems for studying many aspects of fungal diseases, ranging from molecular mechanisms of pathogenesis and host cultivar/species specificity to fungal population dynamics (KANG et al. 2000; VALENT 1990). The power of the rice blast system has been further enhanced with the availability of genome sequence of both M. grisea and rice. The goal of this study is to answer the question of how AVR genes of M. grisea have evolved. Understanding the evolution and mechanisms of variation of AVR genes in M. grisea will provide insights into designing durable resistance in rice.

The research described in this thesis aims to provide a comprehensive understanding of the mechanisms underpinning the molecular evolution of the MgMEP AVR gene family in M. grisea. Data in this thesis show that the MgMEP gene family consists of at least three family members and its evolution may involve functional divergence. The telomeric locations and diverse DNA repeat sequences are likely responsible for the dynamic nature of some of the gene family members. Data from this study led to the following three hypotheses to be tested in the future: (1) members of the MgMEP gene family function as host-specific virulence factors; (2) the evolution of the MgMEP gene family may involve modification of protein to preserve peptidase activity while avoiding recognition as avirulence factor; (3) members of the MgMEP gene family have evolved through gene duplication, deletion, and rearrangement via recombination events mediated by DNA repeats flanking the gene family.

The first hypothesis is supported by structural conservation of the MgMEP genes in M. grisea isolates from diverse hosts as well as previous studies suggesting infection-specific induction of the O-137 MgMEPI gene expression and its putative peptidase
activity which might modify proteins in the host plant (Bryan et al. 2000; Jia et al. 2000; Orbacher et al. 2000). The absence of the gene family in some M. grisea isolates suggests that their function is not essential, but may confer a strain or lineage-specific function. It needs to be tested whether MgMEP genes encode functional metallopeptidases and if the peptidase activity is critical for any virulence activity.

The evolution of AVR genes may involve functional divergence separating avirulence and putative virulence functions. This is supported by the observations that many MgMEP genes lack avirulence activity but maintain the structure as putative metallopeptidases, suggesting that the evolution of the MgMEP genes may involve modifications of gene products to preserve peptidase activity while avoiding recognition as avirulence factor. In particular, the MgMEP3 gene, a recently duplicated gene, lacks avirulence activity but appears to encode a potentially functional metallopeptidase. One possible scenario for explaining MgMEP3 gene evolution would be that after gene duplication, the MgMEP3 gene may have mutated to alter its recognition as an avirulence factor, while preserving peptidase activity. Such an evolutionary mechanism would challenge the prediction that an AVR gene that imposes higher fitness penalties for its loss-of-function will rarely mutate, thus the R gene that targets such AVR gene will be more durable.

Comparative analysis of genome organization of different members of the MgMEP gene family provides clues to the genetic instability of certain members. Genetically unstable MgMEP1/MgMEP2 genes that show frequent amplification/deletions are located near telomeres and are flanked by DNA repeats. This result supports the hypothesis that the presence of AVR genes in the highly dynamic telomere regions may provide an advantage to the pathogen by allowing it to evade the R gene-mediate recognition rapidly. Further sequence analysis of the MgMEP gene family-associated telomeres and their flanking regions in isolates from diverse hosts will be needed to test this hypothesis.
A targeted gene replacement method (termed ATMT-DS) described in this thesis will facilitate the functional analysis of fungal genes of interest. Members of a gene family may exhibit functional redundancy. Determining the function of such genes, including members of the MgMEP gene family, requires cumulative disruption of all the genes. ATMT-DS will be useful for efficiently generating such mutants. ATMT-DS is designed to enrich the gene targeted mutants rather than increasing the efficiency of homologous recombination. Better understanding of the molecular mechanisms underpinning homologous recombination will facilitate the improvement of ATMT-DS.
Appendix

TRANSFORMATION AND MUTAGENESIS OF FILAMENTOUS FUNGI

Chang Hyun Khang¹, Sook-Young Park¹, Hee-Sool Rho², Yong-Hwan Lee², and Seogchan Kang¹
¹Department of Plant Pathology, The Pennsylvania State University, University Park, PA 16802, USA
²School of Agricultural Biotechnology, Seoul National University, Seoul 151-742, Korea

Note: This appendix has been submitted for publication.

Abstract

A better understanding of fungal biology will facilitate judicious use of beneficial fungi and will also advance our efforts to control pathogenic fungi. Molecular studies of fungal biology have been greatly aided by transformation-mediated mutagenesis techniques. Transformation via non-homologous integration of plasmid DNA bearing a selectable marker (e.g., antibiotic resistance gene) has been widely used for the random insertional mutagenesis of fungi, as an alternative to chemical and radiation mutagens, mainly because the integration into the genome of plasmid provides a convenient tag for subsequent identification and isolation of the mutated gene. Homologous recombination between a target gene on the chromosome and the introduced DNA carrying its mutant allele results in targeted gene knock-out. An important advance in fungal transformation methodology is the development of Agrobacterium tumefaciens-mediated transformation (ATMT) protocols for fungi. ATMT has been successfully applied to a phylogenetically diverse group of fungi and offers a number of advantages over conventional transformation techniques in both the random insertional mutagenesis and targeted gene knock-out. In this chapter, we describe ATMT protocols and vectors for fungal gene
manipulation using two plant pathogenic fungi, *Magnaporthe grisea* and *Fusarium oxysporum* as target organisms

1. Introduction

Transformation-mediated forward and reverse genetic analyses have greatly facilitated functional studies of fungal genes (KANG and DOBINSON 2004). In most filamentous fungi, transformation results from the integration of introduced DNA into the fungal genome by either non-homologous or homologous recombination. Mutagenesis of fungi through random integration of the transforming DNA via non-homologous recombination has been widely used for tagging and isolating fungal genes that are involved in various aspects of fungal biology (KANG and METZENBERG 1993; LINNEMANNSTÖNS et al. 1999; LU et al. 1994; SWEIGARD et al. 1998). Mutagenesis of a specific gene via homologous recombination between the target gene and its mutant allele that has been introduced through transformation is also routinely performed in many fungi. A typical procedure for transforming filamentous fungi involves the preparation of fungal protoplasts, delivery of the transforming DNA with associated selectable marker, and selection of the generated transformants. The protoplast generation step, which involves the digestion of the fungal cell walls using a mixture of hydrolytic enzymes, is often critical for high transformation efficiency and is difficult to optimize for reproducibly producing good protoplasts. *Agrobacterium tumefaciens*-mediated transformation (ATMT) offers an alternative and versatile means for transforming fungi without relying on protoplasts.

tumefaciens appears to utilize the same virulence genes that are required for plant transformation (BUNDOCK et al. 1995; MICHELSE et al. 2004a), suggesting that the bacterium utilizes a conserved mechanism for transformation regardless of hosts. In addition to the ability to transform spores, hyphae, and even mushroom fruiting body tissue, ATMT exhibited a number of other advantages as a method for forward and reverse genetic analyses of fungi: (i) high transformation efficiency, resulting in several hundred transformants per $1 \times 10^6$ spores in many fungal species (DE GROOT et al. 1998; MULLINS et al. 2001; RHO et al. 2001) (ii) increased frequency of homologous recombination, a feature conducive for efficient targeted gene knock-out (BUNDOCK et al. 1999; DOBINSON et al. 2004; ZWIERS and DE WAARD 2001), and (iii) low copy number of inserted T-DNA per genome (less than two on average), which facilitates the identification of a gene tagged by the T-DNA especially in asexual fungi (DE GROOT et al. 1998; MULLINS et al. 2001; RHO et al. 2001; SULLIVAN et al. 2002).

Although targeted gene knock-out via transformation is feasible in many fungi, for fungi that exhibit low frequencies of homologous recombination, identification of the desired mutant often requires that a large number of transformants be generated and screened (GARDINER and HOWLETT 2004; KHANG et al. 2005). To circumvent this time-consuming process, a new targeted gene knock-out method, which is based on ATMT with a mutant allele of the target gene flanked by a conditional negative selection marker, was developed (see Fig. 7). A dual (positive and negative) selection of transformants permitted the enrichment of target mutants (GARDINER and HOWLETT 2004; KHANG et al. 2005). This method, termed ATMT-DS, can potentially be applied to a broad spectrum of fungi and may serve as a powerful functional genomic tool in fungal research. In this chapter, protocols and vectors for ATMT and targeted mutagenesis of filamentous fungi are described using two plant pathogenic fungi, **M. grisea** and **Fusarium oxysporum**, as target organisms. In both fungi, we typically obtain 100-400 transformants per $1 \times 10^6$ spores. These tools can also be applied to other fungi with no or little modification.
2. Materials

2.1. Fungal and Bacterial Strains (see Note 1)

4. *Agrobacterium tumefaciens* strains AGL1 and EHA105 (Klee 2000).
5. *Escherichia coli* strain XL1-Blue MRF’ (Stratagene, La Jolla, CA).

2.2. Culture media

All culture media use 1.5% agar (Difco, Sparks, MD) to solidify medium and are autoclaved (120°C for 15 min) before use.

1. Oatmeal agar: After incubating 50 g of rolled oats (Quaker Oats, Chicago, IL) in 500 mL at 70°C for one hour, filter them through cheesecloth. Add agar and bring the volume to one liter.
2. Complete medium: 6 g of yeast extract, 6 g of casein acid hydrolysate, 10 g of sucrose per liter.
3. 2YEG: 2 g of yeast extract and 10 g of glucose per liter.
4. Carnation leaf agar: After sterilizing 1.5% agar in water by autoclaving, cool it down to 60°C. Pour agar solution in Petri plates (90 mm in diameter) containing sterilized (via irradiation) carnation leaf pieces (20-30 pieces per plate; see Note 2).
5. Potato dextrose: 24 g of potato dextrose (Difco, Sparks, MD) per liter.
6. Aspergillus minimal medium: For one liter of medium, add 50 mL stock salt solution (60 g NaNO₃, 5.2 g KCl, 5.2 g MgSO₄·7H₂O and 8.2 g KH₂PO₄ per liter), 10 g Glucose, and 2 mL Hutner’s trace elements (2.2 g ZnSO₄·7H₂O, 1.1 g H₂BO₃, 0.5 g MnCl₂·4H₂O, 0.5 g FeSO₄·7H₂O, 0.16 g CoCl₂·6H₂O, 0.16 g CuSO₄·5H₂O, 0.11 g (NH₄)₆Mo₇O₂₄·4H₂O and 5 g Na₂EDTA per 100 mL).

7. CMC broth: 15 g of carboxymethyl cellulose, 1 g of yeast extract, 0.5 g of MgSO₄, 1 g of NH₄NO₃, and 1 g of KH₂PO₄ per liter (see Note 3).

8. Water agar: Autoclave 1.5% agar in water.

9. LB: 10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl per liter.

10. Minimal medium (MM): See Table 6 and Note 4.

11. Induction medium (IM): See Table 6 and Note 5.

12. Co-cultivation medium (CM): See Table 6 and Note 5.

### Table 6. Preparation of Stock Solutions and Transformation Media

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Stock solution (100 mL)a</th>
<th>MMb</th>
<th>IMc</th>
<th>CMd</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-buffer (pH 7.0)</td>
<td>K₂HPO₄</td>
<td>20 g</td>
<td>1 mL</td>
<td>1 mL</td>
</tr>
<tr>
<td></td>
<td>KH₂PO₄</td>
<td>14.5 g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-N</td>
<td>MgSO₄·7H₂O</td>
<td>3 g</td>
<td>2 mL</td>
<td>2 mL</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>1.5 g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1% CaCl₂·2H₂O</td>
<td>CaCl₂·2H₂O</td>
<td>1 g</td>
<td>0.1 mL</td>
<td>0.1 mL</td>
</tr>
<tr>
<td>Spore elements</td>
<td>ZnSO₄·7H₂O</td>
<td>0.01 g</td>
<td>1 mL</td>
<td>1 mL</td>
</tr>
<tr>
<td></td>
<td>CuSO₄·5H₂O</td>
<td>0.01 g</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H₂BO₃</td>
<td>0.01 g</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSO₄·H₂O</td>
<td>0.01 g</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Na₂MoO₄·2H₂O</td>
<td>0.01 g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20% NH₄NO₃</td>
<td>NH₄NO₃</td>
<td>20 g</td>
<td>0.25 mL</td>
<td>0.25 mL</td>
</tr>
<tr>
<td>20% Glucose</td>
<td>Glucose</td>
<td>20 g</td>
<td>1 mL</td>
<td>1 mL</td>
</tr>
<tr>
<td>0.01% FeSO₄</td>
<td>FeSO₄</td>
<td>0.01 g</td>
<td>1 mL</td>
<td>1 mL</td>
</tr>
<tr>
<td>50% Glycerol</td>
<td>Glycerol</td>
<td>50 mL</td>
<td>-</td>
<td>1 mL</td>
</tr>
<tr>
<td>1M MES (pH 5.3)</td>
<td>MES</td>
<td>21.32 g</td>
<td>-</td>
<td>4 mL</td>
</tr>
<tr>
<td>Sterile H₂O</td>
<td>-</td>
<td>93.5 mL</td>
<td>88.5 mL</td>
<td>88.5 mL</td>
</tr>
<tr>
<td>Kanamycin stock</td>
<td>-</td>
<td>0.15 mL</td>
<td>0.15 mL</td>
<td>0.15 mL</td>
</tr>
<tr>
<td>Acetosyringone</td>
<td>-</td>
<td>0.2 mL</td>
<td>0.2 mL</td>
<td></td>
</tr>
<tr>
<td>Agar</td>
<td>-</td>
<td>-</td>
<td>1.5 g</td>
<td></td>
</tr>
</tbody>
</table>

*a* Prepare all the stock solutions (except acetosyringone) in deionized H₂O as indicated. It is not necessary to sterilize the acetosyringone stock. Glucose, FeSO₄, kanamycin, and MES are sterilized by filtration through a 0.22-micron filter, and the remaining reagents are sterilized by autoclaving for 20 minutes at 120°C. Store them at 4°C except FeSO₄, kanamycin, acetosyringone, and MES (stored at -20°C).

*b* Minimal medium (see Note 4).

*c* Induction medium (see Note 5).

*d* Co-cultivation medium. CM is identical to IM except the addition of 1.5% agar.
2.3. Antibiotics and Other Chemical Agents (see Note 6)

1. Carbenicillin: Prepare as a stock of 100 mg/mL in H₂O.
2. Kanamycin: Prepare as a stock of 50 mg/mL in H₂O.
3. Rifampicin: Prepare as a stock of 10 mg/mL in methanol.
4. Acetosyringone (AS, 3,5-dimethoxy-4-hydroxyacetophenone): Prepare as a stock of 200 mM in 95% ethanol (see Note 7).
5. Hygromycin B (hygB): Prepare as a stock of 100 mg/mL in H₂O.
6. Cefotaxime, Sodium salt: Prepare as a stock of 200 mM in H₂O.
7. Moxalactam, Sodium salt: Prepare as a stock of 100 mg/mL in H₂O.
8. F2dU (5-fluoro-2’-deoxyuridine): Prepare as a stock of 10 mM in H₂O.
9. MES (1 M): dissolve 42.64 g of MES in 160 mL of deionized H₂O. The pH is adjusted to 5.3 with 10 N NaOH with stirring vigorously on a magnetic stirrer until the MES has completely dissolved. Then, deionized H₂O is added to bring the final volume to 200 mL (see Note 8).

2.4. Binary Vectors for Fungal Transformation (see Note 9 and Fig. 12)

1. pBHt2: A vector carrying the hpt (hygB phosphotransferase) gene under the control of the Aspergillus nidulans TrpC promoter on the T-DNA.
2. pKHt: A vector carrying the hpt gene under the control of the A. nidulans TrpC promoter plus the ColE1 replication of origin and the chloramphenicol resistance gene on the T-DNA.
3. pDHt: A vector for the targeted mutagenesis of fungal genes via double recombination.
4. pGKO1: A vector for the targeted mutagenesis of fungal genes via ATMT-DS.
5. pGKO2: A vector for the targeted mutagenesis of fungal genes via ATMT-DS.
Figure 12. Schematic diagrams of the T-DNA region in five binary vectors for fungal transformation. The LB and RB of the T-DNA are denoted by vertical lines. The orientation of transcription from \( hpt, \) \( cam \) (conferring resistance to chloramphenicol) and \( HSVtk \) is indicated by arrow (5' to 3'). ColE1 ori indicates the ColE1 replication origin. MCS1 corresponds to the multiple cloning site of pCAMBIA1300 (offering nine unique restriction sites in pBHt2 and pDHt). In pGKO1, due to the presence of one or more of the \( \text{Bam}H\text{I}, \text{Pst}I, \text{Sac}I, \text{Sal}I, \) and \( \text{Sma}I \) site in \( HSVtk \), only four restriction sites in MCS1 can be used for cloning a mutant allele. The multiple cloning site of pGKO2 (designated as MCS-KS) was derived from pGreenII0000 (HELLENS et al. 2000) and offers 15 unique restriction sites. The \( \text{Bam}H\text{I}, \text{Eco}R\text{V}, \text{Pst}I, \text{Sac}I, \text{Sal}I, \) and \( \text{Sma}I \) sites that are present in \( HSVtk \) were removed via site-directed mutagenesis to produce \( HSVtk \) (M). Drawings are not to scale.

2.5. Other Solutions and Supplies

1. Cryostorage tubes (2 mL).
2. 80% sterilized glycerol.
3. Sterilized soda lime glass beads (3 mm in diameter; see Note 10).
4. Hemacytometer.
5. Cheesecloth.
6. Nitrocellulose membrane (47 mm in diameter).
7. Pasteur pipet (melt the tip in flame to produce a round bulb).
8. Genomic DNA extraction buffer: 50 mM Tris-HCl, pH 7.5, 100 mM EDTA, 0.5% SDS, 300 mM Sodium Acetate, pH 6.
9. Proteinase K (10 mg/mL)
10. TE saturated phenol:chloroform (1:1).
11. Isopropanol.
12. 70% ethanol.
13. RNase A (10 mg/mL).
14. 1X TE buffer: 10 mM Tris-HCl, pH 7.5 and 1 mM EDTA.
15. Ethidium bromide: Prepare as a stock of 10 mg/mL in 1X TE.
16. DNA labeling kit.
17. 0.2 mL thin-wall PCR tubes
18. Taq DNA polymerase and reaction buffer

3. Methods

3.1. Transformation of *A. tumefaciens* with Binary Vector

3.1.1. Preparation of Competent Cells

1. Streak *Agrobacterium* cells stored in 20% glycerol stock at -80°C on LB agar plate amended with appropriate antibiotic (30 µg/mL of rifampicin for strain EHA105 and 100 µg/mL of carbenicillin for strain AGL1).
2. Incubate the plate at 25-28°C until colonies appear (~2 days).
3. Inoculate a single colony into 4 mL LB containing appropriate antibiotic and grow at 28°C, 250 rpm to log phase (OD\textsubscript{600} = 0.3 – 0.6).

4. Inoculate the 4 mL culture into 100 mL LB in a 500 mL flask and grow at 28°C, 250 rpm to OD\textsubscript{600} = 0.5.

5. After chilling the culture for 10 minutes on ice, pellet \textit{Agrobacterium} cells by centrifugation at 3,000 \( g \) for 10 minutes at 4°C.

6. Remove LB and resuspend the cells with 2 mL of 20 mM CaCl\textsubscript{2}.

7. After dispensing cells into 1.5 ml microcentrifuge tubes (100-200 µL per tube), freeze them in liquid nitrogen and store at –80°C until needed.

3.1.2. Transformation (see \textbf{Note 11})

1. Add 1 µg of vector DNA (in about 5 µl) in 1.5 mL microcentrifuge tube.

2. Thaw competent cells on ice.

3. Dispense 50 µl of competent cells into each microcentrifuge tube containing DNA and mix them by gently pipetting up and down 2-3 times.

4. Freeze in liquid nitrogen for 5 minutes (see \textbf{Note 12}).

5. Heat shock the frozen cells at 37°C for 20 minutes.

6. After adding 0.7 ml of LB, grow cells at 28°C, 250 rpm for 2 hours.

7. Spread cells on LB agar plate containing 75 µg/mL kanamycin (for all the vectors described in Section 2.4) using glass beads (see \textbf{Note 10}) and incubate the plates (upside down) at 28°C until transformant colonies appear (1.5-2 days).

8. Pick two independent transformants and culture them in MM amended with 75 µg/mL kanamycin (see \textbf{Note 13}).
3.2. Transformation of *M. grisea* and *F. oxysporum*

3.2.1. Preparation of *M. grisea* Spores

1. Inoculate *M. grisea* on oatmeal agar and grow for 1-2 weeks under constant fluorescent light at room temperature.
2. Harvest *M. grisea* spores by scraping fungal culture with 1 mL micropipet tip after flooding the plate with 10 mL of sterile water.
3. Filter the spore suspension through two layers of cheesecloth to remove large debris.
4. Determine spore concentration using a hemacytometer. Adjust the spore concentration to 10⁶ spores/mL (see Note 14). If the original solution is too diluted, concentrate them by resuspending with a smaller volume of water after centrifugation.

3.2.2. Preparation of *F. oxysporum* Spores

1. Inoculate *F. oxysporum* spores to CMC broth and culture on a rotary shaker (100 rpm) at room temperature for a week (see Note 15).
2. Filter the culture through two layers of cheesecloth to remove mycelia.
3. Harvest spores by centrifugation at 3000g for 5 minutes followed by two washes with sterile water.
4. Resuspend spores in sterile water and adjust the spore concentration to 10⁶ spores/mL (see Note 14).
3.2.3. Transformation

1. Incubate the *Agrobacterium* strain containing a binary vector described in Section 2.4 in 1 mL MM amended with 75 µg/mL kanamycin at 28°C, 250 rpm for 2 days.
2. The *Agrobacterium* cells are diluted to OD<sub>600</sub> = 0.15 in IM amended with kanamycin and 200 µM acetosyringone (AS) and incubate additional 6 hours at 28°C, 250 rpm (OD<sub>600</sub> reaches around 0.6).
3. Mix 100 µL of fungal spore suspension with 100 µL of *Agrobacterium* cells in a microcentrifuge tube and spread the mixture on a nitrocellulose membrane placed on CM in a small Petri plate (15x50 mm) (see Note 16).
4. Incubate the plate for 2 days at 25°C in dark (see Note 17).
5. Transfer the membrane onto appropriate selection medium amended with hygB for selecting fungal transformants. For *M. grisea*, use complete medium agar amended with 250 µg/mL hygB, 200 µM cefotaxime, and 100 µg/mL moxalactam. *Aspergillus* minimal medium amended with 75 µg/mL hygB, 200 µM cefotaxime, and 100 µg/mL moxalactam is used for selecting *F. oxysporum* transformants (see Note 18).
6. Incubate the plate at 25°C until hygB-resistant colonies appear (see Note 19).

3.2.4. Isolation and Purification of Transformants

1. Prepare 24-well microtiter plates that contain appropriate medium for sporulation: oatmeal agar for *M. grisea*, and CMC broth for *F. oxysporum* (see Note 20).
2. Transfer a little bit of mycelia from individual transformants to microtiter well using a sterile toothpick or a fine tipped forcep and incubate the plate at 25°C (see Note 21).
3. For *M. grisea*, flood each well with sterile water and pipet up and down to dislodge conidia. *F. oxysporum* culture in CMC requires no treatment.
4. After plating conidia on solid medium (water agar for *M. grisea* and potato dextrose agar for *F. oxysporum*) amended with appropriate antibiotics and incubating for 24 hours, pick one or two single spores for each transformant using a microscope and transfer them on appropriate culture medium for subsequent analysis and preservation (see Note 22).

3.2.5. Molecular Analysis of Transformants

Southern analysis, PCR, or both have been applied for: (i) testing the intactness of a gene (e.g., a reporter gene such as GFP) introduced via ATMT, (ii) determining the copy number of inserted T-DNA, (iii) identifying gene knock-out mutants, and (iv) isolating genomic regions flanking inserted T-DNA.

3.2.5.1. Southern Analysis (see Note 23)

1. Prepare 24-well plates containing 1 mL of liquid medium (2YEG for *M. grisea* and potato dextrose broth for *F. oxysporum*). Given the mitotic stability of inserted T-DNA (Mullins et al. 2001; Rho et al. 2001), after purifying transformants through single spore isolation, it is not necessary to add antibiotics in growth medium.

2. Inoculate individual transformants to be analyzed in each well and seal the plate with parafilm to prevent evaporation of medium.

3. Grow the transformants for a week at room temperature (RT) on a rotary shaker set at 100 rpm.

4. Grind the cultures in 24-well plate using a hand-made 24-well format grinder (see Note 24).

5. After adding 0.6 mL extraction buffer and 0.5 µL proteinase K per well, seal the 24-well plate with aluminum foil tape and mix thoroughly by inverting the plate.
6. Incubate at 65°C for 1 hour.
7. Transfer the samples into individual microcentrifuge tubes and add 0.6 mL TE saturated phenol:chloroform (1:1). Mix well by inverting the tubes several times.
8. Centrifuge for 10 minutes at 12,000g at RT to separate the two phases.
9. Transfer the aqueous layer to clean tubes and precipitate DNA by adding 420 µL isopropanol (0.7 volume of the sample) and mixing the content.
10. Centrifuge for 10 minutes at 12,000g at RT to pellet DNA.
11. After discarding supernatant, wash DNA with 70% ethanol and dry.
12. Resuspend DNA in 20-40 µL TE. Dissolve pellet by gently tapping the tubes.
13. Treat dissolved DNA with 1 µL of RNase A (10 mg/mL) for 30 minutes at 37°C.
14. Digest 10 µL of DNA with appropriate restriction enzyme for 2-3 hours in a total volume of 30 µL (see Note 25).
15. Run digested DNA on 0.7% agarose gel (see Note 26) and blot the gel on a hybridization membrane.
16. Hybridize the blot with appropriate probe.

3.2.5.2. Identification of Gene Knock-Out (KO) Mutants by PCR (see Note 27)

1. Design a pair of primers, one corresponding to the 3’- or 5’-end of the positive selection marker (hygB resistance gene) and the other corresponding to a target gene. Design another pair of primers specific to a different gene in transformed fungus for testing the quality of DNA template (see Note 28).
2. Run PCR under appropriate conditions (e.g., primer annealing temperature, extension time according the expected size of PCR product, etc.).
3. Analyze PCR products by running them on agarose gel. Expected patterns are illustrated in Fig. 13.
3.2.5.3. Thermal Asymmetric Inter-Laced (TAIL) PCR

Random, insertional mutagenesis of the fungal genome via ATMT is an efficient forward genetic tool for isolating mutants that are defective in various aspects of fungal biology. Inserted T-DNA provides a molecular tag for isolating genes mutagenized by insertion. TAIL-PCR can be used to isolate the genomic DNA of fungal transformant flanking the site of T-DNA insertion. This method allowed efficient isolation of the host DNA sequence flanking a site of T-DNA insertion in *F. oxysporum* (MULLINS et al. 2001). Plasmid rescue can also be used as an alternative method to TAIL-PCR (see Note 29).

1. Synthesize primers binding T-DNA border sequence as well as one or more arbitrary degenerate (AD) primer as shown in Fig. 14 (see Note 30).
2. Run PCR reactions as previously described (MULLINS et al. 2001).
3. Separate PCR products on 1% agarose gel to isolate appropriate fragments (see Note 31).
4. Sequence the isolated fragments using the outermost specific primer, either RB-3 or LB-3, as the sequencing primer.

**Figure 14.** Primers for TAIL-PCR. Right (A) and left (B) border ends of the T-DNA sequence showing the position and melting temperature ($T_m$) of the primers employed for TAIL-PCR. Bold uppercase letters indicate the 24 bp imperfect repeat (LB and RB), which is cleaved at the indicated positions prior to insertion. The sequence and $T_m$ of the degenerate arbitrary primer (AD-1) is also shown (C). In AD-1, $W = A$ or $T$ and $N = A, C, G, T$. 
3.3. Targeted Gene Knock-Out (KO) via ATMT-DS

3.3.1 Transformation

1. Transform fungal spores as described above (see Section 3.2), with Agrobacterium cells containing a gene KO vector that carries a mutant allele created by inserting the hpt gene flanked by the tk gene (a conditional negative selection marker conferring sensitivity to F2dU) under the control of a fungal promoter (see Fig. 7).

2. HygB-resistant transformants are subjected to negative selection by transferring individual transformants using sterilized toothpicks to a selection medium containing 5 µM F2dU in addition to hygB and cefotaxime (see Note 32).

3. Incubate the transformants for 3-5 days at 25°C in dark.

4. Transfer viable transformants to appropriate medium in 24-well microtiter plates for sporulation.

5. Isolate single spores as described above (see Section 3.2.4) and analyze purified transformants for the presence of desired mutation (see Note 33).

4. Notes

1. To ensure phenotypic stability of fungal strains, avoid subculturing them. Repeated subculturing on nutrient-rich medium can potentially lead to loss of virulence and fertility and reduced conidiation (FISHER et al. 1982; VALENT et al. 1986). Thus, it is highly recommended that a fresh culture should be activated from a stock stored at a non-metabolizing state prior to experiment. For long-term storage of F. oxysporum, conidia are stored in 15% glycerol at -80°C or fungal cultures grown on carnation leaf agar (see Section 2.2) are preserved as described in Fisher et al. (FISHER et al. 1982) at -20°C. Stock cultures of M. grisea are prepared and stored as
described by Valent et al. (Valent et al. 1986). After placing sterilized Whatman filter paper disks on oatmeal agar or complete medium plate (see Section 2.2), fungal culture is allowed to grow over the disks. Once the fungus has completely colonized the disks, the disks are removed and dried. The disks are placed in sterilized post stamp envelope or 2 ml cryo-storage tube and stored at -20°C.

2. Sterilized carnation leaf pieces can be purchased from the Fusarium Research Center (FRC) at Penn State (http://www.cas.psu.edu/Docs/CASDept/Plant/FRC/; 814-863-0145).

3. It can take several hours for CMC to dissolve completely. While vigorously stirring the solution, add CMC small portion at a time so that it does not form a big clump. Heating shortens the time needed for dissolving CMC.

4. MM (sans FeSO₄ and an appropriate antibiotic) can be stored at 4°C for several months. Appropriate amount of 0.01 % FeSO₄ and an antibiotic are added into MM prior to growing Agrobacterium strains.

5. IM (sans FeSO₄, MES, an appropriate antibiotic, and acetosyringone) can be stored at 4°C for several months. For preparing CM, after autoclaving and cooling down the medium to around 70°C, add 0.01% FeSO₄, 1M MES, and acetosyringone as described in Table 1.

6. Prepare antibiotics stock solutions using deionized water and sterilize them by filtering through a 0.22 µm filter. All the antibiotics stock solutions, except hygB (4°C), are stored at -20°C.

7. Acetosyringone stock solution forms crystals at -20°C. These crystals will quickly dissolve in ethanol when the tube is inverted several times (or briefly vortexed) at room temperature.

8. The filter-sterilized MES stock is stored frozen at –20°C. When the frozen solution is thawed, precipitation appears as white powder. Vortex vigorously until the precipitate has completely dissolved.

9. The vectors described here were derived from pCAMBIA1300 (http://www.cambia.org/main/r_et_camvec.htm). Construction scheme for these vectors was previously described (Khang et al. 2005; Mullins et al. 2001).
10. These beads are quite handy for quickly and uniformly spreading bacterial/fungal cells on solid medium, especially when a large number of samples need to be plated (3-4 beads per 90 mm plate). Wash the beads by soaking them in 1 N HCl for several hours. Rinse them with water thoroughly until pH is above 5, dry in a baking oven and autoclave.

11. Since only one or two *A. tumefaciens* transformants are needed for subsequent step, scaling down the protocol (1/2 to 1/4) often works fine.


13. If not used immediately, store these cultures at –80°C until needed as 20% glycerol stock. After fungal transformation, save the one that has yielded higher transformation efficiency as a permanent stock for future use. If one wants to determine whether gene(s) on the introduced binary vector is intact, PCR with primers specific to the target gene can be performed using 1-2 µL of intact cells in MM as a template. Alternatively, one can extract plasmid DNA from 5 mL bacterial culture using a commercial silica membrane-based spin column and perform restriction enzyme analysis. Although the yield of plasmid DNA is much lower than that from *E. coli*, we find it sufficient for restriction enzyme analysis.

14. Optimal spore concentration may vary depending on fungal strain/species. The efficiency of transformation via ATMT varied significantly (more than two orders of magnitude) among several fungal species (De Groot *et al.* 1998). Even within a species, we noticed significant strain-dependent variation in transformation efficiency (Mullins *et al.* 2001). Therefore, determining transformation efficiency for a target strain/species is highly recommended prior to launching a large-scale transformation. The *M. grisea* and *F. oxysporum* strains used here (see Section 2.1) typically yield 10~40 transformants per membrane (100-400 transformants/1x10^6).

15. Alternatively, culture *F. oxysporum* on carnation leaf agar, which takes much longer than culturing in CMC to produce enough spores for transformation.

16. Use of nitrocellulose membrane does not appear to be essential for successful transformation. The number of plates needed varies depending on the objective of transformation. For generating a pool of insertional mutants randomly tagged by the
T-DNA, a large number of plates are needed. However, for introducing a reporter gene (such as the GFP gene) to a fungal strain, one plate should be sufficient.

17. Duration of cocultivation is one of the parameters that may require optimization for efficiently transforming a new fungal strain/species. Although transformation efficiency generally increases with the duration of cocultivation in both *F. oxysporum* and *M. grisea* (Mullins *et al.* 2001; Rho *et al.* 2001), excessive mycelial growth during prolonged co-cultivation might make it difficult to subsequently identify individual transformants. For certain fungal species such as *Fusarium graminearum*, mycelial growth is rapid and excessive during cocultivation, which appears to inhibit bacterial growth and leads to very low transformation efficiency.

18. Geneticin can also be used for selecting transformants from both species (when transformed with a vector carrying the gene conferring resistance to this antibiotic). Optimal hygB (and geneticin) concentration for selecting transformants may vary significantly from one species to another. Sensitivity to these antibiotics can also vary depending on media used. Prior to transforming a new fungal species via ATMT, one must determine the degree of antibiotic sensitivity of the species by inoculating untransformed strain on several media amended with varying concentrations of chosen antibiotic (24 well microtiter plate is suitable for this testing). Use a concentration that completely blocks the growth of untransformed strain for selecting transformants.

19. Bacterial cells frequently form a thick lawn on the membrane during two days of cocultivation. Reducing the amount of bacterial cells from the membrane seems to increase the transformation efficiency and speed up the growth of transformants. Place the membrane upside down on selection medium (90 mm Petri plate) for a few minutes. Drag around the membrane on the surface of the medium to remove bacterial cells as much as possible, and subsequently flip and place the membrane on the same medium. Transformant colonies mainly form on the membrane with colonies appearing occasionally outside of the membrane. It takes approximately 4-7 days to observe transformant colonies.
20. In addition to hygB (or geneticin), add cefotaxime and moxalactam to the medium to ensure that no *Agrobacterium* cells transferred along with fungal transformants can grow.

21. It may take up to a week to produce conidia from *M. grisea*, whereas 2-3 days of culture is often sufficient to produce conidia from *F. oxysporum*.

22. Mix a small volume of conidial suspension (1-3 µL) with a sterilized water drop (100-200 µL) on the surface of medium in small Petri plate (50 mm in diameter). Spread conidia using glass beads (see Note 10). For isolating single spores from a large number of transformants, we use the following method: (i) Prepare a thin layer of medium inside the lid of regular Petri plate (90 mm in diameter), (ii) Draw lines on the plastic surface to divide the medium into 16 sections (number them), (iii) Touch the culture of transformant with the rounded tip of Pasteur pipet (produced by melting the tip in flame), (iv) Streak spores attached to the tip on the medium as if plating bacterial cells, (v) After sterilizing the pipet tip, plate another transformant to the next section, (vi) Cover the medium with the bottom part of Petri plate and seal with parafilm, (vii) Grow at 25°C for 24 hours, and (viii) Cut out an agar block containing a single germinating conidium using the flattened tip of a pin (or a platinum wire) under a compound microscope (100X magnification).

23. This DNA extraction protocol is for quickly analyzing transformants and may not be suitable for analyses requiring a high quality genomic DNA preparation. DNA yield from *F. oxysporum* is significantly lower than that from *M. grisea*, mainly due to less mycelial growth under this condition. For more genomic DNA, culture *F. oxysporum* transformants in potato dextrose broth in small Petri plates (after inoculation, leave them on lab bench for a week).

24. It is not essential to remove the medium. Alternatively, transfer the cultures into individual microcentrifuge tubes after blotting between paper towels and grind them using a Teflon bar. Freeze-drying is not necessary, but doing so increases DNA yield.

25. For determining the copy number of inserted T-DNA, use a restriction enzyme that has no or one recognition site on the T-DNA constructed used for transformation.
For checking gene disruption, use a restriction enzyme that has no recognition site on the positive selection marker (e.g., hygB-resistance gene) interrupting the target gene.

26. Before running the gel, run 5 µL of digested DNA in 0.7% agarose minigel to confirm complete digestion. If DNA has not been completely digested, further purify genomic DNA: After increasing the volume of DNA solution to 200 µL with TE, repeat a phenol/chloroform extraction. For DNA precipitation, add 10 µL of 4 M NaCl and 400 µL of 100% ethanol. Resuspend DNA in TE.

27. PCR analysis can be performed prior to single spore isolation to reduce the number of transformants that need to be purified. A protocol by Xu and Hamer (XU and HAMER 1995) is suitable for quickly analyzing transformants by PCR without purifying genomic DNA.

28. The target gene-specific primer should anneal to the outside of the gene fragment used for creating a mutant allele (see Fig. 13) so that when combined with a primer specific to the positive selection marker, only gene KO mutants yield a PCR product. The second pair of primers should produce a PCR product from all transformants and serve to test the quality of PCR template. Alternatively, instead of using two pairs of PCR primers, a single pair of primers that bind to the outside of the mutant allele can be used for both purposes. Ectopic transformants should produce a PCR product that corresponds to the wild type gene, whereas gene KO transformants produce a PCR product that is larger than that from ectopic transformants due to the presence of the positive selection marker. Although this method requires less PCR reactions than the former method, the size of both PCR products to be amplified is larger than that by the former method, thus requiring a longer PCR extension time. Because it also amplifies two PCR products that differ by the size of the positive selection marker (e.g., 1.4 kb for the hpt gene), optimizing PCR conditions might be more difficult.

29. If the border sequences become truncated beyond the annealing site for the amplifying borders, PCR amplification of the desired product would not occur. Although the frequency of T-DNA truncation in F. oxysporum and M. grisea does
not appear to be as high as that in plants, such events have been detected. Unless the truncation is extensive, reaching deeply into the T-DNA, plasmid rescue can solve this problem. The binary vector pKHt (see Section 2.1) carries the ColE1 replication of origin and the chloramphenicol resistance gene (Mullins et al. 2001). Therefore, from mutants generated using pKHt, inserted T-DNA along with its flanking genomic DNA can be cloned into *E. coli* as a replicating plasmid as follows: (i) Digest genomic DNA of selected transformant with a restriction enzyme that does not have a recognition site on the T-DNA, (ii) Ligate digested DNA, and (iii) Transform *E. coli* and isolate chloramphenicol resistant colonies. The fungal genes tagged by T-DNA of pKHt can also be identified via TAIL-PCR.

30. We designed a 16-nucleotide (nt) long arbitrary degenerate (AD) primer, labeled AD-1, with a calculated melting temperature (T_m) of 47°C. T_m’s of primers specific for each border sequence of the T-DNA (LB-1, 2 and 3 for the left border and RB-1, 2 and 3 for the right border), ranging from 20 to 26 nt in size, were designed to be at least 58°C or higher, as previously recommended (Liu and Whittier 1995). The 5’-end of LB-3 and RB-3 was 74 bp and 63 bp apart from the 5’-end of LB-2 and RB-2, respectively, to facilitate the identification of PCR products corresponding to T-DNA insert junctions by size comparison between the secondary and tertiary reaction products.

31. As a result of the nested arrangement of the RB and LB primers within one another (see Fig. 14), the size profile of the desired PCR product is staggered, decreasing with each TAIL-PCR reaction. The primary reaction typically produces several products. The number of PCR products is reduced with each successive reaction (see Note 30). In some transformants, more than one PCR product can be amplified from a single insertion site, presumably due to the presence of multiple annealing sites of AD-1 in the vicinity of the inserted T-DNA.

32. Ectopic transformants expressing the *tk* gene fail to grow on this medium due to their sensitive to F2dU, whereas gene KO mutants are insensitive to F2dU (see Fig. 7). Add 100 µL of selection medium in each well of 96-well microtiter plate. Alternatively, prepare the medium in Petri plates and spot transformants in a grid
format. The concentration of F2dU required for selecting against ectopic transformants varied significantly from one species to another (KHANG et al. 2005). The medium used can also affect sensitivity to F2dU. For a new fungal species, determine the degree of its F2dU sensitivity by comparing the growth of tk transformants and untransformed strain on several media amended with varying F2dU concentrations (use 24 well microtiter plate for this testing). Use a concentration that completely blocks the growth of transformants expressing tk.

33. Certain fraction of transformants after dual selection appears to be false positive (i.e., ectopic transformants insensitive to F2dU) in both M. grisea and F. oxysporum. This type of transformants mainly results from truncation of the tk gene during T-DNA integration (KHANG et al. 2005). The frequency of false positives varies depending on a number of factors, including fungal species and strains, the target locus, and the length of the gene fragments used for constructing a mutant allele. In fungal species/strains that have a high rate of false positive, PCR prior to single spore isolation (see Note 27) is highly recommended.

Acknowledgements

The methods described here were developed using grants from USDA-NRI (2002-02367), the Crop Functional Genomics Center of the 21st Century Frontier Research Program (CG1413), and the Korea Science and Engineering Foundation (Agricultural Plant Stress Research Center). We would like to thank Justin Dillon for critically reviewing the chapter.
BIBLIOGRAPHY


BENNETZEN, J. L., and E. A. KELLOG, 1997 Do plants have one-way tickets to genomic obesity? Plant Cell **9**: 1509-1514.


BROSIOUS, J., 1999 RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. Gene 238: 115-134.


Plant Cell 8: 1747-1755.

DABOUSSI, M. J., 1996 Fungal transposable elements: generator of diversity and genetic 

DANGL, J. L., and J. D. G. JONES, 2001 Plant pathogens and integrated defence responses 
to infection. Nature 411: 826-833.


Research Foundation, Silver Spring, MD.

DE GROOT, M. J. A., P. BUNDOCK, P. J. J. HOYKAAS and A. G. M. BEIJERSBERGEN, 
Nature Biotech. 16: 839-842.

DESLANDES, L., J. OLIVIER, N. PEETERS, D. X. FENG, M. KHOUNLOTHAM et al., 2003 
Physical interaction between RRS1-R, a protein conferring resistance to bacterial 
wilt, and PopP2, a type III effector targeted to the plant nucleus. PNAS 100: 
8024-8029.

Mapping of avirulence genes in the rice blast fungus, Magnaporthe grisea, with 

DIXON, R., 1986 The phytoalexin response: Elicitation, signalling, and control of host 

DOBINSON, K. F., S. J. GRANT and S. KANG, 2004 Cloning and targeted disruption, via 
Agrobacterium tumefaciens-mediated transformation, of a trypsin protease gene 

DOBINSON, K. F., R. E. HARRIS and J. E. HAMER, 1993 Grasshopper, a long terminal 
repeat (LTR) retroelement in the phytopathogenic fungus Magnaporthe grisea. 
Molecular Plant Microbe Interaction 6: 114-126.

DODDS, P. N., G. J. LAWRENCE, A.-M. CATANZARITI, M. A. AYLIFE and J. G. ELLIS, 
2004 The Melampsora lini AvrL567 avirulence genes are expressed in haustoria 
and their products are recognized inside plant cells. Plant Cell: tpc.020040.


IKEDA, K.-I., H. NAKAYASHIKI, T. KATAOKA, H. TAMBA, Y. HASHIMOTO et al., 2002 Repeat-induced point mutation (RIP) in *Magnaporthe grisea*: implications for its sexual cycle in the natural field context. Molecular Microbiology **45**: 1355-1364.


KEEN, N. T., S. TAMAKI, D. KOBAYASHI, D. GERHOLD, M. STAYTON *et al.*, 1990 Bacteria expressing avirulence gene D produce a specific elicitor of the soybean hypersensitive reaction. Molecular Plant Microbe Interaction **3**: 112-121.


MEHDY, M. C., 1994 Active oxygen species in plant defense against pathogens. Plant Physiology **105**: 467-472.


MONTENEGRO-CHAMORRO, M. V., 1997 Allelic diversity of an avirulence genes in Colombian field isolates of the rice blast fungus, pp. Purdue University, West Lafayette.


MUNDT, C. C., 1990 Probability of mutation to multiple virulence and durability of resistance gene pyramids. Phytopathology **80**: 221-223.


ROHE, M., A. GIERLICH, H. HERMANN, M. HAHN, B. SCHMIDT et al., 1995 The race-specific elicitor, NIP1, from the barley pathogen, Rhynchosporium secalis, determines avirulence on host plants of the Rrs1 resistance genotype. EMBO Journal 14: 4168-4177.


SCHECHTMAN, M. G., 1990 Characterization of telomere DNA from *Nerospora crassa*. Gene **88**: 159-165.


STAKMAN, E., 1915 Relation between *Puccinia graminis* and plants highly resistant to its attack. Journal of Agricultural Research **4**: 193–199.


EDUCATION

Present  Ph.D. Candidate Plant Pathology, Penn State University, University Park, PA
02/00  M.S. Plant Pathology, Seoul National University, Seoul, Korea
08/97  B.S. Agricultural Biology, Seoul National University, Seoul, Korea

TEACHING EXPERIENCE

Fall 2003  Teaching Assistant, Penn State University, Microbe-Plant Interactions Lab (PPATH 405)
Summer 2003  Research Mentor, Penn State University, PA Governor’s School Independent Study

HONORS AND AWARDS

The Genetics Society of America Financial Aid Award, 2005
The Penn State Research Foundation Inventor Incentive Award, 2004
College of Agricultural Sciences Graduate Student Travel Award, 2004
Dr. Henry W. Popp Plant Pathology Travel Award, 2004

PUBLICATIONS


PATENTS


PUBLISHED ABSTRACTS


PROFESSIONAL MEMBERSHIP

The American Phytopathological Society (APS) 2002-Present