2013

CHARACTERIZATION AND DISTRIBUTION OF NOVEL NON-LTR RETROELEMENTS DRIVING HIGH TELOMERE RFLP DIVERSITY IN CLONAL LINES OF MAGNAPORTHE ORYZAE

John H. Starnes
University of Kentucky, jhstar0@uky.edu

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

John H. Starnes, Student

Dr. Mark L. Farman, Major Professor

Dr. Lisa J. Vaillancourt, Director of Graduate Studies

CHARACTERIZATION AND DISTRIBUTION OF NOVEL NON-LTR RETROELEMENTS DRIVING HIGH TELOMERE RFLP DIVERSITY IN CLONAL LINES OF *MAGNAPORTHE ORYZAE*

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Agriculture
at the University of Kentucky

By
John Howard Starnes

Lexington, KY

Director: Dr. Mark L. Farman, Professor of Plant Pathology

2013

ABSTRACT OF DISSERTATION

CHARACTERIZATION AND DISTRIBUTION OF NOVEL NON-LTR RETROELEMENTS DRIVING HIGH TELOMERE RFLP DIVERSITY IN CLONAL LINES OF *MAGNAPORTHE ORYZAE*

The filamentous ascomycete fungus *Magnaporthe oryza*e is a pathogen of over 50 genera of grasses. Two important diseases it can cause are gray leaf spot in *Lolium perenne* (perennial ryegrass) and blast in *Oryza sativa* (rice). The telomeres of *M. oryzae* isolates causing gray leaf spot are highly variable, and can spontaneously change during fungal culture. In this dissertation, it is shown that a rice-infecting isolate is much more stable at the telomeres than an isolate from gray leaf spot. To determine the molecular basis of telomere instability several gray leaf spot isolates telomeres were cloned, which revealed two non-LTR retrotransposons inserted into the telomere repeats. The elements have been termed *Magnaporthe oryzae* Telomeric Retrotransposons (MoTeRs). These elements do not have poly-A tails common to many other non-LTR retrotransposons, but instead have telomere like sequences at their 5' end that allow them to insert into telomeres. Intact copies of MoTeRs were restricted to the telomeres of isolates causing gray leaf spot. Surveys for the presence of these elements in *M. oryzae* showed they were present in several host-specialized forms including gray leaf spot isolates, but were largely absent in the rice blast isolates. The absence of MoTeRs in rice blast isolates, which are relatively stable by comparison, suggested that the telomere instability in gray leaf spot isolates could be due to MoTeRs. Analyzing spontaneous alterations in telomere restriction fragment profiles of asexual progeny revealed that MoTeRs were involved. Expansion and contraction of MoTeR arrays were observed and account for some telomere restriction profile changes. New telomere formation in asexual progeny followed by MoTeR addition was also observed. Based on this evidence, MoTeRs are largely responsible for the high variability of telomere restriction profiles observed in GLS isolates.

KEYWORDS: Non-LTR, retrotransposon, telomere, gray leaf spot, *Magnaporthe oryzae*

John Howard Starnes
Student's Signature

February 06, 2013
Date

CHARACTERIZATION AND DISTRIBUTION OF NOVEL NON-LTR RETROELEMENTS DRIVING HIGH TELOMERE RFLP DIVERSITY IN CLONAL LINES OF *MAGNAPORTHE ORYZAE*

By

John Howard Starnes

Dr. Mark L. Farman
Director of Dissertation

Dr. Lisa J. Vaillancourt
Director of Graduate Studies

February 6th, 2013

ACKNOWLEDGEMENTS

The production of this dissertation would not have been possible without the support and guidance of many wonderful individuals. I want to thank my family for their unfailing support in my endeavors.

The assistance and training by current and past members of the Farman laboratory was invaluable. My advisor, Dr. Mark Farman, has been instrumental in facilitating my professional development. He has provided me with greater insight into molecular genetics, plant pathology, and scientific thinking. Dr. David Thornbury was a boundless resource in the laboratory. He provided technical assistance and advice whenever needed. Melanie Heist provided an extra pair of ears, eyes, and hands in any complications that I experienced while completing my work.

The members of my committee (Dr. Christopher Schardl, Dr. Paul Vincelli, and Dr. Randal Voss) have provided needed support and direction during my pursuits, and were very patient in my progress.

I was able to pursue my research effectively with funding from the Department of Plant Pathology and the Research Challenge Trust Fund Fellowship. Funding for the research described in this dissertation was provided by the National Science Foundation.

TABLE OF CONTENTS

Acknowledgement..... iii

List of Tables..... iv

List of Figures..... viii

CHAPTER ONE: Introduction and Literature Review..... 1
1.1 Introduction..... 1
1.1 Host-specialization of *M. oryzae*..... 4
1.2 Biology and population dynamics of *M. oryzae*..... 7
1.3 Transposable elements..... 11
1.3.1 Distribution of Class I transposable elements in *M. oryzae*... 15
1.3.2 Distribution of Class II transposable elements in *M. oryzae*. 19
1.4 Implications of TEs in the variability of *M. oryzae*..... 20
1.5 Telomere variability in *M. oryzae*..... 22
1.6 Variability of GLS pathogens and the presence of MoTeRs..... 27
1.7 Overview of dissertation research..... 29

CHAPTER TWO: Materials and Methods..... 34
2.1 Fungal cultures..... 34
2.2 Plant inoculation and conidia collection..... 34
2.3 Small scale DNA extraction..... 35
2.4 Large scale DNA preparation..... 36
2.5 End-enriched clone library preparation..... 36
2.6 Cloning of MoTeR Bands unlinked to telomeres..... 38
2.7 Primer design..... 39
2.8 Hybridization probes..... 39
2.9 Southern hybridization..... 40
2.10 Cloning of MoTeR-to-MoTeR junctions..... 42
2.11 Sequencing..... 42
2.12 Tandem repeat analyses..... 43
2.13 Relatedness of MoTeR1 to other retrotransposons..... 43
2.14 Statistical methods in RFLP comparisons..... 44
2.15 Phylogenetic analyses of MoTeR1 and gene loci..... 44

CHAPTER THREE: Characterization and Variation of MoTeR Repeats within GLS Isolates..... 56
3.1 Introduction..... 56
3.2 Results..... 59
3.2.1 MoTeRs structure and classification..... 59
3.2.2 Genomic distribution of MoTeRs among GLS isolates..... 63
3.2.3 Variability of MoTeRs among Lolium pathotype isolates.... 66
3.2.4 Structural Variability within FH..... 69
3.2.5 Factors involved in size variation of MoTeRs..... 70

3.3 Discussion..... 74
3.3.1 MoTeRs are members of a new class of non-LTR retrotransposons..... 74
3.3.2 Variability of the 5' end of MoTeRs..... 76
3.3.3 Possible endonuclease site of MoTeRs..... 78
3.3.4 MoTeRs distribution..... 80
3.3.5 Telomere instability in subculture..... 81

CHAPTER FOUR: Generation of New Telomeric Profiles in Mitotic Progeny..... 97
4.1 Introduction..... 97
4.2 Results..... 99
4.2.1 Rates of telomere change in a GLS isolate versus a RB isolate..... 99
4.2.2 Telomere fragment variation in mitotic progeny of other GLS isolates..... 102
4.2.3 Comparison of telomere and Pot2 profile variation..... 103
4.3 Discussion..... 105
4.3.1 Instability of telomere restriction fragments in mitotic progeny of GLS isolates..... 105
4.3.2 Comparative instability of telomere and Pot2 restriction fragments..... 108
CHAPTER FIVE: Molecular Basis for Telomere Variability in LpKY97-1A..... 116
5.1 Introduction..... 116
5.2 Results..... 117
5.2.1 MoTeRs association with *de novo* telomere profiles in LpKY97-1A..... 117
5.2.2 Methylation in telomere profile changes..... 118
5.2.3 Characterization of de novo telomere fragments..... 119
5.2.3.1 rDNA telomere truncations..... 122
5.2.3.2 Capture and duplication of internal sequence at a telomere..... 123
5.2.3.3 Truncation of MoTeR2 arrays..... 125
5.3 Discussion..... 130
5.3.1 MoTeRs in telomere restriction fragment length polymorphisms..... 130
5.3.2 rDNA rearrangements..... 131
5.3.3 Duplication and capture of internal sequence at telomeres..... 134
5.3.4 MoTeR transposition..... 135
5.3.5 Molecular mechanisms underlying MoTeR array contraction and extension..... 137

CHAPTER SIX: Distribution and Evolutionary History of *Magnaporthe oryzae* Telomeric Retrotransposons with Magnaporthe 150
6.1 Introduction 150
6.2 Results 152
6.2.1 Distribution of MoTeRs within Magnaporthe 152
6.2.2 MoTeR evolution within *Magnaporthe oryzae* 154
6.2.3 Possible mechanisms affecting the evolutionary history of MoTeRs 158
6.3 Discussion 159
6.3.1 Distribution of MoTeRs in Magnaporthe 159
6.3.2 Possible routes to the present evolutionary pattern of MoTeR1 161

CHAPTER SEVEN: Conclusions 176
7.1 MoTeRs are active retrotransposons in the isolates from the Lolium pathotype 176
7.2 Possible role of MoTeRs in increasing the phenotypic variability 177
7.3 Implications of research to disease management 179

Appendices
Appendix A: Sequences of MoTeR1, MoTeR2, and NhTeR1 181
Appendix B: MoTeR Expression Preliminary Experiment 186
B.1 Introduction 186
B.2 Methods 186
B.3 Results and Discussion 187
Appendix C: LpKY97-1A and IT11 Alignments and Analyses of MoTeR1 190

References 192

Vita 225

LIST OF TABLES

Table 1-1 Transposable elements in Magnaporthe..... 32
Table 2-1 Magnaporthe isolates used in the experiments..... 45
Table 2-2 Primers used in the experiments..... 50
Table 2-3 Reverse transcriptases used to determine the class of *M. oryzae* retrotransposons..... 52
Table 3-1 Short Tandem Repeats within *Magnaporthe oryzae* Telomeric Retrotransposons..... 83
Table 3-2 MoTeR cohybridization with telomeric fragments in GLS isolates..... 84
Table 3-3 Copy number analyses of telomere and MoTeRs using densitometric scanning and visual estimation..... 85
Table 3-4 5' end variation in MoTeR-to-MoTeR junctions..... 86
Table 4-1 Telomeric *Pst*I restriction fragment changes of three generations of mitotic progeny in LpKY97-1A..... 110
Table 4-2 Telomeric *Pst*I restriction fragment changes in three GLS isolates and one rice infecting isolate..... 111
Table 4-3 Changes in EcoRI restriction fragments with Pot2 probes and telomere probes..... 112
Table 5-1 MoTeR associated telomere restriction fragment changes in mitotic progeny of LpKY97-1A..... 141
Table 6-1 Hybridization signals of MoTeR probes in various host pathotypes of Magnaporthe..... 167
Table 6-2 Telomeric association of MoTeR hybridizing fragments within host pathotypes of Magnaporthe..... 168
Table 6-3 Maximum composite likelihood estimate of the pattern of nucleotide substitution..... 169

LIST OF FIGURES

Figure 1-1 Comparison of homologous chromosome end structure in FH and 70-15 33
Figure 3-1 Structure of *Magnaporthe oryzae* Telomeric Retrotransposons (MoTeRs) 87
Figure 3-2 Phylogenetic relationship of MoTeR1 to other reverse Transcriptases 88
Figure 3-3 Comparison of the restriction enzyme-like (REL) endonuclease domains of MoTeR1 with closely related non-LTR retrotransposons 89
Figure 3-4 Southern blot analyses of genomic DNA from GLS isolates of *M. oryzae* 90
Figure 3-5 Internal MoTeR1 fragments cloned from PstI digests of genomic DNA from RGNJ and TFRGA 91
Figure 3-6 Expected composition of hybridization signals in the Southern hybridizations of *Bam*HI digests probed with MoTeR(5') in the gray leaf spot isolate FH 92
Figure 3-7 Southern analyses of genomic DNA from various GLS isolates 93
Figure 3-8 Graphical representation of the three different MoTeR1 elements in the FH isolate 94
Figure 3-9 Characterization of size variation of MoTeRs in GLS isolates 95
Figure 4-1 Southern analyses comparing stability of telomeric DNA fingerprints in 70-15 and LpKY97-1A 113
Figure 4-2 Telomere fingerprint stability of mitotic progeny from KS320 and RGNJ, which are Lolium pathotype isolates of *M. oryzae* 114
Figure 4-3 Southern analyses of Pot2 probe and telomeric probes in 70-15 and LpKY97-1A 115
Figure 5-1 Southern analyses of LpKY97-1A mitotic progeny using telomere, MoTeR1, and MoTeR2 probes 142
Figure 5-2 Methylation in telomere restriction fragment length polymorphisms 143
Figure 5-3 Characterization of telomere profile variation in mitotic progeny in LpKY97-1A 144
Figure 5-4 Characterized telomeres in LpKY97-1A 145
Figure 5-5 Southern analyses to detect changes at the 31B chromosome end in mitotic progeny of LpKY97-1A 146
Figure 5-6 Resolution of 31B fragments in Southern analyses using either *Apa*I or *Pst*I digested genomic DNA 147
Figure 5-7 Southern analyses of progeny from three third generation variants of telomere 31B 148
Figure 5-8 Southern analyses of progeny from four third generation variants of telomere 31B 149
Figure 6-1 Southern analyses of genomic DNA of *M. oryzae* isolates from Triticum pathotype isolates 170

Figure 6-2 Southern analyses of genomic DNA from Setaria pathotype isolates of *M. oryzae*..... 171
Figure 6-3 Southern analyses of genomic DNA from Oryza pathotype isolates of *M. oryzae*..... 172
Figure 6-4 Southern analyses of genomic DNA from various pathotypes of Magnaporthe..... 173
Figure 6-5 Southern analyses of genomic DNA from Digitaria pathotype isolates of *Magnaporthe grisea*..... 174
Figure 6-6 Neighbor-joining (NJ) phylogenetic trees based on concatenated sequence from three genomic loci [MPG1, CH7-Bac7, and the internal transcribed spacer (ITS)] (A) and MoTeR1 (B) in isolates from various *Magnaporthe oryzae* pathotypes..... 175

CHAPTER ONE

Introduction and Literature Review

1.1 Introduction

Rice (*Oryza sativa*) is an important food staple worldwide with 650 billion tons produced annually. To increase the production of rice, the diseases of rice that reduce yield need to be controlled. The most important disease is rice blast (OU 1980), which is caused by the filamentous ascomycete fungus *Magnaporthe oryzae* (anamorph *Pyricularia oryzae*). The characteristic symptoms of rice blast (RB) disease are gray to white ellipsoid lesions and lodging of the panicle, which results in failure to produce seed. Significant yield losses are common in rice fields infected with *M. oryzae* (BONMAN *et al.* 1991; REDDY and BONMAN 1987).

M. oryzae is not limited to causing disease in rice. It can infect over 50 monocot species (URASHIMA *et al.* 1999). These include other economically important crops such as wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), and several different millets. *M. oryzae* can infect all above ground parts of the plant (OU 1980), and was described as being able to infect roots as well (DUFRESNE and OSBOURN 2001). Pathogen activity can result in a variety of symptoms, which include blast of the inflorescence as described above, leaf spots, and foliar blight.

Gray leaf spot (GLS), another disease caused by *M. oryzae*, has emerged as a problem in major turfgrass species including tall fescue (*Lolium arundinaceum* Darbyshire), perennial ryegrass (*Lolium perenne*), and annual ryegrass (*Lolium multiflorium*)(UDDIN *et al.* 2003). Isolates of *M. oryzae* that infect these grasses are

morphologically indistinguishable (COUCH and KOHN 2002; ROSSMAN *et al.* 1990; YAEGASHI and UDAGAWA 1978), but evolutionarily distinct from the *M. grisea* isolates that cause GLS disease in crabgrass (Digitaria sp.) *(COUCH and KOHN 2002)*.

The GLS disease caused by *M. oryzae* was first noticed in 1957 on St. Augustinegrass (*Stenotaphrum secundatum*) in Florida (MALCA and OWEN 1957). GLS symptoms on St. Augustinegrass start as small brown to red spots and expand rapidly into larger leaf spots. The leaf spots are tan in color when dry and gray when wet. Borders of the round to oblong leaf spots are brown to red in color with a chlorotic halo around some of the outside edges. Leaf spots may coalesce into larger lesions (MALCA and OWEN 1957). GLS can further cause stolons and leaves to die leading to a scorched appearance of the turf (ATILANO and BUSEY 1983; FREEMAN 1962).

In the 1970s, outbreaks of GLS were observed in annual ryegrass. The disease severity was much worse in annual ryegrass than what had been previously observed in the St. Augustinegrass and resembled symptoms more similar to blast disease, such as foliar blighting and leaf spot (BAIN *et al.* 1972).

Starting in the early 1990s, GLS was reported on perennial ryegrass in golf course fairways (LANDSCHOOT and HOYLAND 1992). The damage from outbreaks of GLS in perennial ryegrass was extensive. Sporadic outbreaks that have occurred since then have resulted in a significant loss of turf (DERNOEDEN 1996; UDDIN *et al.* 2003). Symptoms of GLS start as small water-soaked leaf spots, which further expand into gray to light brown necrotic spots. The borders of the leaf spots are purple to dark brown in color often with chlorotic halos. Foliar blighting can occur

when leaf spots coalesce. As the disease progresses complete necrosis of the leaves and death of the plant may occur (UDDIN *et al.* 2003).

GLS disease can be controlled through three principal control measures: cultural practices, fungicides, and cultivar resistance (OU 1980). There are problems with each of these control measures when used alone. Cultural practices do not provide a sufficient amount of control when disease pressures are high (VINCELLI 2000). Fungicides are available for disease control, but they must be applied as a preventative. This represents an expense that many turfgrass managers cannot justify, when an outbreak of disease is uncertain. Additionally, *M. oryzae* populations can gain resistance to fungicides when used intensively (VINCELLI and DIXON 2002).

Development of GLS resistant cultivars would alleviate the reliance on fungicides, and would be a valuable control measure. Many of the current cultivars are susceptible to *M. oryzae* (BONOS *et al.* 2005; HOFMANN and HABLIN 2000). Potential sources of resistance to GLS have been identified (BONOS *et al.* 2004; FRASER 1996; PEYYALA and FARMAN 2006), and some are currently being developed in commercial cultivars (BONOS *et al.* 2005). The long-term efficacy of disease resistant cultivars has been relatively low because of the extreme variability of *M. oryzae*. Races of the fungus can overcome plant resistance, resulting in new epidemics (ZEIGLER and CORREA 2000). Typically, because of the issues with individual control practices, outlined above, an integrated program is needed to provide better control of the disease.

The generation of new variants of the fungus is an important mechanism in defeating the effectiveness of at least two of the control measures, fungicides and cultivar selection. A better understanding of possible sources of new variability could lead to better control of the fungus. Thus, the purpose of this dissertation was to identify some of the possible mechanisms by which new variability in *M. oryzae* populations might arise in GLS isolates.

1.2 Host-specialization of *M. oryzae*

M. oryzae, as a species, can infect over 50 species of grasses, but individual strains of the pathogen are limited to certain hosts (BORROMEO *et al.* 1993; DOBINSON *et al.* 1993; MACKILL and BONMAN 1986). The individual strains are grouped into host-specialized forms, or pathotypes, based on the host from which they were collected. Within each of the host-specialized forms some individuals will cause disease in some varieties of the host plant but not in other varieties. These individuals are further grouped into races based on the host plant varieties in which they infect.

Two types of genetic resistance of grass hosts to *M. oryzae* have been identified. The resistance can be controlled by a large number of genes with small individual effects (quantitative), or one to few genes having large effects (qualitative) (AHN and OU 1982; TALBOT and FOSTER 2001). Resistance to RB and GLS has been attributed to both quantitative and qualitative resistance (AHN and OU 1982; BONOS *et al.* 2005; ZEIGLER and CORREA 2000).

Plants with quantitative resistance, also called partial resistance or horizontal resistance, are susceptible to infection, but typically the severity of disease is reduced (CHEN *et al.* 2003). This type of resistance reduces the growth

and reproduction of a pathogen, which can slow the rate of epidemic development (AHN and OU 1982; ASHIZAWA *et al.* 1999; CASTANO *et al.* 1989). Partial resistance can limit the generation of variability in the pathogen by decreasing the number of disease cycles in a season (generations in which to evolve pathogenicity) and by reducing sporulation.

A single to a few genes typically controls qualitative resistance, also called vertical or specific resistance. Qualitative resistance typically inhibits the initial establishment of the pathogen. Interactions between *M. oryzae* and its host plants under qualitative resistance appear to be incompatible. For example, the host may seem to be immune, produce a hypersensitive reaction, or slow the reproduction of *M. oryzae*. The plant genes conferring qualitative resistance are known as resistance (R) genes (FLOR 1971). Plants containing the R gene are resistant to *M. oryzae* isolates carrying a corresponding gene known as an avirulence gene (*Avr* gene) (FLOR 1971; JIA *et al.* 2000). The plant resistance is a result of the activation of host defense responses, such as hypersensitive response (HR). Avr genes may encode effector proteins that suppress basal defense responses (PARK *et al.* 2012), and losses of Avr genes in other fungal species have led to reduced virulence (BOLTON and TOMMA 2008; VAN DEN ACKERVEKEN *et al.* 1993). However, the actual mechanism by which most Avr gene products in *M. oryzae* interact with their host are unknown (STERGIOPOULOS and DE WIT 2009).

Avr genes can function in determining the cultivar specificity, and they can also control the host specificity of different *M. oryzae* isolates (KOBAYASHI *et al.* 1989). One example of host specificity control is a gene family that confers

avirulence in a strain of *M. oryzae* to weeping lovegrass (*Eragrostis curvula*). One member of this family, *PWL2*, is located internally on the chromosome, and is spontaneously lost in mutant progeny of an avirulent parental isolate. *PWL2* based resistance was stable in a rice strain Guy11, but this was due to having two copies of the gene in different chromosomal locations. Segregation of these genes through genetic crosses revealed that the two *PWL2* genes (*PWL2-1* and *PWL2-2*) were both unstable in the progeny, with frequent, spontaneous, virulent mutants being produced. When both genes are absent the progeny can infect weeping lovegrass (SWEIGARD *et al.* 1995). Though *PWL2* is not located at the chromosome end, it is located near transposable elements (DEAN *et al.* 2005). The stability of the *Avr* genes may be affected by the genetic elements that surround them (KANG and YONG-HWAN 2000). Another example is Avr-C039, from a weeping lovegrass pathogen (FARMAN and LEONG 1998), which confers avirulence of *M. oryzae* to rice cultivars containing the corresponding R gene Pi-CO39(t) (CHAUHAN *et al.* 2002). Functional copies of AVR-CO39 are present in other host-specific pathotypes (PEYYALA and FARMAN 2006; TOSA *et al.* 2005), but largely absent in Oryza pathotype (FARMAN *et al.* 2002; TOSA *et al.* 2005). There is currently a disagreement in the literature over whether Avr-CO39 is a species-wise or cultivar-wise host specificity gene (TOSA *et al.* 2005; ZHENG *et al.* 2011).

Loss of *Avr* gene function or absence of the *Avr* gene in an isolate of *M. oryzae* allows the pathogen to avoid detection by host plants containing the corresponding R gene. The loss of the Avr gene would change the race of that pathogen. Novel virulent variants are known to frequently arise during rice cultivation (BONMAN

1992). In large natural populations of some fungi, deletion of avirulence genes is a common evolutionary mechanism in gaining virulence (GOUT *et al.* 2007; SCHÜRCH *et al.* 2004; WOUW *et al.* 2010). This represents a source of variability that *M. oryzae* can exploit to circumvent defense mechanisms in the host, and considering the highly clonal population structure of the fungus these mechanisms may be important in race dynamics. For example, studies of the avirulence gene *Avr-Pita* have shown that spontaneous mutants of an avirulent *M. oryzae* parental isolate, that have lost *Avr-Pita* are virulent on the rice cultivar Yashiro-mochi that carries the R gene *Pi-ta*. The *Avr-Pita* gene is linked to a telomere. Losses of function of the *Avr-Pita* in mutants have been attributed to different factors such as chromosome truncation, mutations in the gene, and a transposable element insertion into the gene (KANG *et al.* 2001; ORBACH *et al.* 2000). Another *Avr* gene linked to a telomere and that mutates frequently is *Avr1-TSUY*. The spontaneous mutants lacking *Avr1-TSUY* can infect cultivar Tsuyuake. Six of these spontaneous mutants were analyzed, and it was revealed that the loss of the gene was due to a deletion mutation. Some of the deletion mutations were believed to be caused by ectopic recombination between homologous transposable elements (KANG and YONG-HWAN 2000). The examples of *Avr-Pita* and *Avr1-TSUY* show that race changes can occur in asexual progeny of a *M. oryzae* isolate.

1.3 Biology and population dynamics of *M. oryzae*

Understanding the basic biology of *M. oryzae* may help to understand where variation might arise within populations. Pathogenic variants within a population may arise through mutation or through sexual recombination. In nature, however,

only the asexual stage has been directly observed. During this stage the somatic hypha grows through the plant and then produces long unbranched conidiophores, which bear groups of three to five bi-septate pyriform conidia. Conidia are produced in a temperature range between 9°C and 35°C, with an optimum of 25°C-28°C. A minimum humidity of 89% is also needed for conidia production. These conidia are then dispersed to potential hosts by water, wind, or some other mechanism (SUZUKI 1975). Once the conidia come in contact with a hydrophobic surface they secrete an apical droplet of spore tip mucilage that attaches the spore to the surface (HOWARD 1994). Conidia germinate if there is free moisture available. An appressorium is produced on the end of the germ tube. A penetration peg then arises from the appressorium, and will use a combination of enzymatic degradation (SKAMNIOTI and GURR 2007) and mechanical force (HOWARD *et al.* 1991) to penetrate the epidermis. It takes six hours at the optimal temperature of 24°C to go from germination of the conidia to the invasion of epidermal cells. A minimum of 12 hours at optimal humidity and temperature is required for optimal disease development. The disease cycle is completed when conidia are produced on the surface of infected tissue after a latent period of five to seven days. When weather conditions are optimal *M. oryzae* progresses through repeated cycles of infection and conidiation leading to disease epidemics (SUZUKI 1975).

Any mutations during asexual propagation could accumulate and lead to variation among different clonal lines. Populations that are older and exclusively clonal should show continuous variation among isolates, ranging from genetically

similar to dissimilar (KUMAR *et al.* 1999). Prolonged asexual reproduction can also lead to an inability of the strain to reproduce sexually (SALEH *et al.* 2012).

M. oryzae's sexual stage has only been visualized *in vitro*. Sexual reproduction in the heterothallic fungus requires the interaction of two mating types, Mat1-1 and Mat1-2 (KANG *et al.* 1994), possibly through a pheromone receptor mediated process (SHEN *et al.* 1999). The sexual stage initiates as a hyphal network of the haploid compatible strains differentiate into female reproductive structures (ascogonia) and the male sexual structures (spermatia released from collars of phialides) (CHUMA *et al.* 2009). Sexual reproduction in ascomycetes occurs when plasmogamy occurs and the nucleus in the spermatia is transferred to the ascogonia through a specialized hyphal structure called the trichogyne. The nuclei remain paired in a dikaryon within the dividing ascogenous hyphae. At the tip of the hyphae the crozier hook will form, and start to develop into the ascus mother cell. The haploid nuclei from strains of opposite mating types fuse to form a diploid zygote. The zygote undergoes meiosis to produce recombinant, haploid ascospores, which are released at maturity from a pore on the top of the flask-shaped perithecium (COPPIN *et al.* 1997).

One important aspect of sexual reproduction is the production of offspring with recombined genotypes due to the independent segregation of unlinked genes during meiosis. The recombined genotypes increase the genotypic diversity which may allow sexual populations to respond more rapidly to selection imposed by resistant cultivars or fungicides (MILGROOM 1996).

There has been some speculation based on population genetic and molecular genetic analyses that sexual recombination is active in *M. oryzae* populations (DOUHAN *et al.* 2011; KUMAR *et al.* 1999). Sexually fertile, hermaphroditic isolates of both mating types have been recovered in natural populations of *M. oryzae* virulent on rice (KUMAR *et al.* 1999). In contrast, all of the *M. oryzae* isolates, which cause GLS on perennial ryegrass, that have been tested are mating-type Mat 1-2 (DOUHAN *et al.* 2011; FARMAN 2002; VIJI and UDDIN 2002). This suggests that sexual reproduction may not play much of a role in the diversity observed in GLS isolates. Sexual reproduction could occur if an isolate virulent on perennial ryegrass were to cross with Mat1-1 isolates from other host pathotypes. However, crosses between different host pathotypes in *M. oryzae* have shown a reduced ability to overcome resistance in adult plants at the level of penetration in the hybrid progeny (MURAKAMI *et al.* 2006). This implies that hybrids may not survive in a natural environment. GLS isolates from perennial ryegrass have been shown to infect other hosts including annual ryegrass, wheat, and weeping lovegrass under artificial conditions (TREDWAY *et al.* 2005; VIJI *et al.* 2001), which suggests that some level of cross-infectivity can occur in nature.

Population structure analyses indicate that GLS isolates specialized to perennial ryegrass (*Lolium perenne*) are genetically similar using a variety of markers including Restriction Fragment Length Polymorphisms (RFLP) (FARMAN 2002), Amplified Fragment Length Polymorphisms (DOUHAN *et al.* 2011; TREDWAY *et al.* 2005), and gene sequences (COUCH *et al.* 2005; VIJI *et al.* 2001). They are most

similar to isolates from other grass hosts including tall fescue, weeping lovegrass, and wheat (TREDWAY *et al.* 2005).

There was some disparity in genetic diversity measures in GLS isolates from perennial ryegrass based on the different probes used in RFLP analyses. Single-copy DNA markers show little variation between isolates. Repetitive transposable element fingerprinting showed higher variation than the single-copy DNA markers, but there was still less than a 15% difference between isolates (FARMAN and KIM 2005). This was similar to AFLP results from GLS isolates collected from tall fescue (TREDWAY *et al.* 2005). A telomere probe showed a different result. It indicated that the chromosome ends of these same perennial ryegrass isolates were highly divergent. Also, mutations that changed telomeric RFLP patterns were readily observed during vegetative growth (FARMAN and KIM 2005).

1.4 Transposable elements

Transposable elements in *M. oryzae* have been used extensively in population studies, and have been linked to disruption of *Avr* genes. The Magnaporthe genome is known to contain many copies of TEs from both major groups. The transposable elements (TE) are divided into two major groups based on the mechanism by which they transpose. In the first class of TEs, a copy of the element transposes via a "copy and paste" mechanism whereby a reverse transcriptase acts upon a RNA intermediate to initiate transposition into a new location within the genome. This class is also known as the retrotransposons. The Class I transposable elements can be further divided into three subclasses: long terminal repeat (LTR) transposable elements contain long terminal repeats and encode a reverse transcriptase; long

interspersed elements (LINEs) lack the long terminal repeats but retain the reverse transcriptase; and short interspersed elements (SINEs) represent short DNA sequences that were reverse-transcribed from RNA and do not have a reverse transcriptase (FINNEGAN 1989; KEMPKEN and KUCK 1998).

The structure and functionality of LTR-transposons are related to retroviruses, except they lack an env gene. The omission of the env gene causes them to be primarily non-infectious. In LTRs, direct sequence repeats flank an internal coding region, which encodes structural and enzymatic proteins. The gag gene encodes the structural proteins. These proteins form the virus-like particle (VLP). The VLP is where the reverse transcription of the LTR's RNA into cDNA will occur. The pol gene encodes proteins, which provide the enzymatic functions of the LTR. The three major proteins encoded by the pol genes are a protease that cleaves the Pol poly protein, the reverse transcriptase (RT) that will be copying the LTR's RNA into cDNA, and the integrase which will integrate the cDNA into a host genome (HAVECKER *et al.* 2004).

Some long interspersed nuclear elements or LINEs have been known to play a role in telomere stability in other organisms (ABAD *et al.* 2004a; TAKAHASHI *et al.* 1997). LINEs are simpler structurally, and more ancient evolutionarily than the LTRs described above. They lack terminal repeats altogether, but some elements may have a poly A-rich sequence at their 3' end. These elements encode a protein with distinct functions: an endonuclease and a reverse transcriptase (CURCIO and DERBYSHIRE 2003). The endonuclease nicks the bottom strand of the target DNA, which results in 3' OH that is used to prime the cDNA synthesis using the LINE RNA

as a template (FENG *et al.* 1996; LUAN *et al.* 1993). Most members of the LINE-like elements are truncated at the 5' end due to incomplete reverse transcription of its RNA template. Currently it is unclear how the cDNA integrates into the upstream end of the target and how second-strand synthesis occurs (CURCIO and DERBYSHIRE 2003). Since the cDNA is incorporated directly into the target DNA in LINEs, unlike the LTR transposons, horizontal transmission is less likely.

The final group of Class I transposable elements are SINEs. They are actively involved in inactivation of genes through transposition (WALLACE *et al.* 1991). The general features of SINEs are an adenine rich 3' end similar to LINE elements, a RNA polymerase III promoter, and a core region (GILBERT and LABUDA 1999). Amplification and dispersion of SINE elements occurs by retrotransposition. The RNA polymerase III-dependent SINE transcripts are reverse transcribed by a LINE element's reverse transcriptase, and the cDNA is then integrated into new genomic sites (JAGADEESWARAN *et al.* 1981; OKADA *et al.* 1997).

In the second class of TEs, the DNA transposes directly from one place to another in the genome using a "cut and paste" mechanism in which the DNA of the transposons is moved to a new location. This class is also known as the inverted repeat DNA transposons (FINNEGAN 1989; WICKER *et al.* 2007). These transposons remove themselves from one place in the genome and insert themselves into a new target with the help of a transposase. The transposase acts by generating free 3' OH ends through hydrolysis of the phosphodiester backbone (excision) at both ends of the transposon. The exposed 3' OH ends will then be inserted into the transposon's target within the genome through a transesterification reaction (insertion). The

insertion of the transposon on each target strand is staggered. As a result, short single-stranded regions, usually two to nine nucleotides in length, will flank the newly inserted transposon. These regions will be repaired by its host's replication machinery causing target site duplications once the segment is repaired (TAVAKOLI and DERBYSHIRE 2001). The gap left by the excised transposon can be repaired by non-homologous end joining or by gene conversion (ENGELS *et al.* 1990; PLASTERK and GROENEN 1992). Inverted repeat transposons expand the genetic diversity of a population not only from the insertion of the elements but also the excision of the elements from the genome.

There are examples of each major class and subclass of TEs in *M.* oryzae (see a description of each element below). TEs are scattered throughout the genome, but may be amplified or clustered in different regions within the genome of *M. oryzae* (REHMEYER *et al.* 2006; THON *et al.* 2004). For example, in the isolate 70-15, TEs make up a higher proportion of the subtelomeres than the genome. The LTR TEs make up approximately 4% of the genome (DEAN *et al.* 2005) and make up approximately 14% of the subtelomeres in the isolate 70-15 (REHMEYER *et al.* 2006). LINEs make up approximately 1% of the genome (DEAN *et al.* 2005) and make up approximately 4% of the subtelomeres (REHMEYER *et al.* 2006). The inverted repeat transposons make up 2% of the genome (DEAN *et al.* 2005) and 5% of the subtelomeres (REHMEYER *et al.* 2006).

The distribution of TEs is varied among different host-specialized isolates, and sometimes distributed differently within a host-specialized group. In the section below the different transposable elements are described and Table 1-1 outlines

some of the distribution differences of TEs among different host-specialized isolates of *M. oryzae*.

1.4.1 Distribution of Class I transposable elements in *M. oryzae*

MAGGY. The most studied long terminal repeat retrotransposon in *M. oryzae* is the Magnaporthe Gypsy-like element (MAGGY), which was first characterized by Farman *et al.* (FARMAN *et al.* 1996b). MAGGY is 5.6 kb in length with 253 bp LTRs that terminate in 6 bp inverted repeats. In isolates containing MAGGY, transposition has been shown to occur during and after the sexual cycle (ETO *et al.* 2001). Most elements are found embedded in AT-rich sequences (FARMAN *et al.* 1996b). The copy number of MAGGY is believed to be controlled by post-transcriptional suppression of MAGGY (NAKAYASHIKI *et al.* 2001a). This element can be activated by stress conditions such as heat shock, copper sulfate, and oxidative stress (IKEDA *et al.* 2001). The distribution of MAGGY in Magnaporthe has been described as limited (ETO *et al.* 2001) to wide (FARMAN *et al.* 1996b). The difference in the classification is largely based on the specific isolates that were used in the studies. MAGGY is present at high copy numbers (>50) in isolates from these pathotypes: Oryza (ETO *et al.* 2001; FARMAN *et al.* 1996b; KUSABA *et al.* 1999), Setaria (ETO *et al.* 2001; FARMAN 2002; FARMAN *et al.* 1996b; KUSABA *et al.* 1999), a Pennisetum (KUSABA *et al.* 1999), and one buffelgrass (*Cenchrus ciliaris*). MAGGY was present at lower copy numbers (<50) in a Bermudagrass (Cynodon dactylon) pathotype isolate (FARMAN *et al.* 1996b). In the Lolium pathotype MAGGY was unevenly distributed (FARMAN 2002; KUSABA *et al.* 1999; TOSA *et al.* 2007). In one study, MAGGY was not observed in the Eragrostis pathotype isolates (ETO *et al.* 2001); while in another study one isolate

had a high copy number (FARMAN *et al.* 1996b). It was believed that the Eragrostis pathotype isolate obtained MAGGY recently in its evolutionary history. This would suggest that it could have arisen through horizontal transmission from another isolate. Horizontal transmission is described as a process by which genetic material is incorporated from one organism to another without being an offspring of that organism. Triticum specialized isolates do not have MAGGY elements (ETO *et al.* 2001) but the MAGGY element can amplify if artificially introduced into their genomes (NAKAYASHIKI *et al.* 2001a).

Grasshopper. The element is 8 kb in length. Grasshopper (*Grh*) contains a smaller LTR (198 bp) than MAGGY. The LTRs in *Grh* terminate in 5 bp inverted repeats, and *Grh* generates target site duplications during its transposition. The distribution of this element is limited to a single subgroup in the Eleusine-specialized isolates (DOBINSON *et al.* 1993; ETO *et al.* 2001), but has also been observed in isolates from *Lolium multiflorum* (KUSABA *et al.* 2006) and the Triticum pathotype.

MGRL-3. The MGLR-3 element was identified and characterized in a mutant of 0-137 that was virulent on the rice cultivar Tsuyuake (KANG 2001). The element is approximately 6.3 kb in length. Protein sequence alignments from the reverse transcriptase (RT) domain indicated it was most similar to the RT proteins encoded by Grasshopper and MAGGY. MGRL-3 was present in all host pathotypes tested including isolates from the Oryza, Triticum, Eleusine, Paspalum, Panicum, Pennisetum, Lolium and Digitaria pathotypes(KANG 2001; VIJI *et al.* 2001). The

highest copy numbers of MGLR-3 are found in the isolates from the Pennisetum, Oryza, and Panicum pathotypes (KANG 2001).

Two additional MGLR-3 like retrotransposons were identified from a Japanese field isolate 9439009, and named *Inago*1 and *Inago*2 (SANCHEZ *et al.* 2011). These retrotransposons appear to be stable within the genome even during induced stress conditions. *Inago*1 was present in high copy numbers in isolates from the Oryza, Setaria, Panicum, Lolium, and Digitaria pathotypes, but had low copy numbers in the isolates from the Eleusine pathotype. *Inago*2 had a similar distribution pattern as *Inago*1 with the exception that there were lower copy numbers in the Digitaria pathotype isolates. This study indicated that the ancestor of these elements was likely present in the genome before the divergence of *M. grisea* and *M. oryzae* (SANCHEZ *et al.* 2011).

Pyret. Another gypsy-like retrotransposon named Pyret is present in the isolates in the Eleusine, Eragrostis, Avena, Triticum, Oryza, and Pennisetum pathotypes. Pyret was absent in Ginger (*Zingiber mioga*) pathotype isolates and in the *Magnaporthe grisea* isolates of the Digitaria pathotype (NAKAYASHIKI *et al.* 2001b).

RETRO6 and RETRO7. RETRO6 and RETRO7 were identified during the genome sequencing of 70-15 (DEAN *et al.* 2005), and there are no studies that address the distribution or structure of these elements. Based on protein sequence alignments these retrotransposons are more similar to the gypsy-like TEs in *M. oryzae*.

RETRO5. RETRO5 is the only LTR retroelement currently identified in *M. oryzae* that is not gypsy-like. It belongs to the Copia class of retrotransposons. The difference between Copia class retrotransposons and Gypsy class transposons is the order in which the proteins are encoded in the sequence. The element is 7.6 kb in length (FARMAN *et al.* 2002). There is a wide distribution of RETRO5 as it is present in *M. oryzae* isolates of the Eleusine, Eragrostis, Stenotaphrum, Triticum, Lolium, Setaria, and Oryza pathotypes. RETRO5 is also found in *M. grisea* isolates of the Digitaria pathotype. These elements are found in the highest copy numbers in Setaria specialized isolates. Currently no studies have addressed the structure of the RETRO5 element.

MGL. The *Magnaporthe grisea* LINE retrotransposon (MGL) is 5.9 kb in length. A small part of the retrotransposon has been used extensively as a probe (MGR583) in RFLP based population studies (DIOH *et al.* 2000; HAMER *et al.* 1989a). MGL was believed to be a LINE-like element due to the lack of long terminal repeats at the 5' and 3' ends of the element (KACHROO *et al.* 1997). There are two open reading frames (ORF) in MGL. The function of the protein encoded by ORF1 is unknown. ORF2 encodes a protein with reverse transcriptase and integrase domains (FUDAL *et al.* 2005). It is present in most host specific isolates, but is found in high copy numbers only in isolates of the Oryza and the Setaria pathotypes (ETO *et al.* 2001).

Mg-SINE. Kachroo et al. identified Mg-SINE as an insertion in the Pot2 transposon (see below). Mg-SINE is a small element only 470 bp in length (KACHROO *et al.* 1995). The 3' end of Mg-SINE is identical to the 3' end of MGL suggesting that it

may use the MGL reverse transcriptase for transposition (THON *et al.* 2004). Mg-SINE is present in isolates of the Oryza, Setaria, Triticum, Eragrostis, Eleusine, Panicum, and Lolium pathotypes (ETO *et al.* 2001).

MINE. The Mixed Interspersed Nuclear Element (MINE) retrotransposon is a 1.9 kb element with a 5' conserved end called WEIRD (bp 1-1114), and a 3' end variable region similar to MGL (bp 1115-1971) including the terminal TAC repeats (FUDAL *et al.* 2005). The MINE element likely uses the reverse transcriptase and integration machinery from the MGL element. High copies of the MINE element were found in isolates of the Pennisetum and Oryza pathotypes. A single copy was present in an Eleusine pathotype isolate, and MINE was absent in a Triticum pathotype isolate. The presence of this transposon in Lolium pathotype isolates is unknown.

1.4.2 Distribution of Class II transposable elements in *M. oryzae*

Pot2. Pot2 was characterized by Kachroo *et. al* in 1994 (KACHROO *et al.* 1994). The element is approximately 1.9 kb in length, and has 43 bp of inverted terminal repeats. Pot2 is present in most host specialized forms of *M. oryzae* (ETO *et al.* 2001). The highest copy numbers are found in the Oryza pathotype isolates.

Pot4. The Pot4 element is similar to Pot2 and was identified during the genome sequencing of 70-15. Not much is currently published that describes Pot4 other than it is approximately 1.9 kb in length (DEAN *et al.* 2005).

Pot3. The transposon is approximately 1.9 kb in length and contains terminal repeats of 42 bp (FARMAN *et al.* 1996a). The MGR586 repeat is part of Pot3 and has been used extensively in population analyses (FARMAN 2002; KUMAR *et al.* 1999; PARK

et al. 2008). This element has a high copy number in Oryza pathotype isolates and their relatives (ETO *et al.* 2001). The Pot3 transposon is apparently absent from Eleusine and Brachiaria pathotypes (FARMAN *et al.* 1996a).

Occan. Occan was discovered by Kito *et al.* 2003 in a spontaneous Pi-a virulent mutant of *M. oryzae*. The element is approximately 2.7 kb in length, and is a member of the Fot1 family of transposons (KITO *et al.* 2003). Occan has a high copy number in isolates of the Oryza pathotype, and has a lower number of copies in isolates of the Setaria and Eleusine pathotypes (KITO *et al.* 2003). No published data is currently available on its distribution in the Lolium pathotype isolates.

1.5 Implications of TEs in the variability of *M. oryzae*

One pattern that emerges with TEs in *M. oryzae* is that the distribution and the copy numbers in host specialized isolates varies in both the location within a genome (XUE *et al.* 2012) and among isolates. The ability of the TEs to influence genome evolution in *M. oryzae* has not been comprehensively explored. There is a correlation between ectopic recombination rate and TE content suggesting that they play a role in generating the genetic diversity observed in *M. oryzae* (THON *et al.* 2006). Accumulation of these elements may lead to chromosome length polymorphisms. TEs may also provide homologous regions for intra-genomic recombination of large sections of the chromosome (DABOUSSI 1997). Ectopic recombination is suggested to have occurred between TEs in the chromosome-unique sequence located adjacent to subtelomeres of *M. oryzae* (REHMEYER *et al.* 2006). The *Avr* genes may be lost or modified during ectopic recombination events between TEs, and if the Avr genes are located near areas of high transposon density

the likelihood that this could occur is increased (FARMAN *et al.* 2002; SWEIGARD *et al.* 1995).

M. oryzae primarily propagates asexually. This increases the odds that mutations or genome rearrangements may become fixed in a population as long as the vegetative fitness and conidia production is not negatively impacted. Genomic rearrangements during asexual propagation may allow populations of *M. oryzae* to change more quickly. TEs can be inactivated during sexual reproduction in other fungi by the Repeat-induced point mutation mechanism (RIP) (CAMBARERI *et al.* 1989; SELKER *et al.* 1987). RIP causes specific mutations from G:C to A:T in repetitive sequences. There is evidence of RIP activity in repetitive elements within the *M. oryzae* genome (DEAN *et al.* 2005; IKEDA *et al.* 2002). Though RIP mutations are typically restricted to repetitive sequences, they have been suggested to "leak" into adjacent single copy sequences (FUDAL *et al.* 2009). In *Neurospora crassa* (IRELAN *et al.* 1994) and *Leptosphaeria maculans* (WOUW *et al.* 2010), for example, RIP mutations were detected in single copy sequences bordering repetitive elements. The leaky nature of RIP could influence the evolution of Avr-genes or other niche associated genes located near repetitive elements in *M. oryzae* during rare sexual reproduction events.

TEs structure and their activity in a genome has profound evolutionary consequences (THON *et al.* 2006). TEs can change the host specificity in the progeny (mitotic or meiotic) through transposition and ectopic recombination. They can disrupt genes or promoters, causing the abolishment of gene function. In one example, a spontaneous mutant of a *M. oryzae* isolate carrying the *Avr-Pita* gene

gained virulence on a rice cultivar containing the Pita R-gene. The virulence gain was due to a Pot3 insertion into the promoter of the *Avr-Pita* gene (ORBACH *et al.* 2000). Virulent isolates have also arisen due to Pot3 transposition into the coding region of *Avr-Pita* gene (ZHOU *et al.* 2007). In another example, a Pot3 insertion into the promoter region of *Avr-Piz-t* in progeny of an avirulent strain has been shown to disrupt proper function, which allowed these isolates to be virulent on rice cultivars containing the corresponding R-gene *Piz-t* (LI *et al.* 2009). Transposition of MGL into a gene controlling conidiophore development led to aberrant conidia morphology which resulted in progeny that produced fewer appressoria and were less pathogenic on plants (NISHIMURA *et al.* 2000). Transposition of MGLR-3 into Pot2, which caused a substantial deletion at the chromosome end, led to virulence on the rice cultivar Tsuyuake in progeny of the *M. oryzae* strain 0-137 (KANG 2001).

These examples indicate that TEs are active and can be attributed as one potential source of the genetic diversity observed in *M. oryzae*, and that TEs can increase or decrease the fitness of a fungal strain. If transposition of TEs allows the fungus to adapt more quickly, *M. oryzae* populations may favor individuals that do not undergo sexual reproduction to ensure that TEs remain active.

1.6 Telomere variability in *M. oryzae*

Interestingly, about half of the characterized *Avr* genes in *M. oryzae* map near chromosome ends (FARMAN 2007). Mutations in telomeric restriction fragments have been observed in Southern hybridization experiments of asexual (FARMAN 2007; FARMAN and KIM 2005) and sexual progeny (CHUMA *et al.* 2011a). This suggests

that chromosome ends might be important in the variability of host-specificity within *M. oryzae*.

M. oryzae, like all eukaryotes, has linear chromosomes, and the extreme ends of the chromosomes are capped by a telomere. Telomeres are made up of tandemly repeated, simple sequences, usually consisting of a GT rich sequence oriented in the 5' to 3' direction towards the chromosome end (ZAKIAN 1995). The telomere sequence found at the end of *M. oryzae* telomeres is (TTAGGG)n (FARMAN and LEONG 1995).

Telomere sequences are bound with multiple proteins, which protect the chromosome ends from degradation (DE LANGE 2005; LOAYZA and DE LANGE 2003; SANDELL and ZAKAIN 1993) and end-to-end fusions (MCCLINTOCK 1939; MCCLINTOCK 1941). The protein complexes that "cap" the telomeres are also known to protect them from homologous recombination and nonhomologous end joining (FERREIRA *et al.* 2004; PALM and DE LANGE 2008). The telomere associated proteins regulate telomere length (KRAUSKOPF and BLACKBURN 1996; VAN STEENSEL and DE LANGE 1997), and are involved in silencing genes located near the telomere (BAUR *et al.* 2001; GOTTSCHLING *et al.* 1990).

In most organisms, simple sequence repeats are maintained by a specialized reverse transcriptase called telomerase. New telomere repeats are added on to the chromosome ends by the enzymatic subunit of telomerase (the telomerase reverse transcriptase-TERT) which uses an "internal" RNA template (TER) to reverse transcribe telomeric sequence at the chromosome end (BLACKBURN and COLLINS

2011; GREIDER and BLACKBURN 1987). Higher telomere instability is observed in mutants which lack telomerase (BLASCO *et al.* 1997; MEYER and BAILIS 2008).

Drosophila melanogaster (the fruit fly) lacks telomerase, and uses a retroelement-mediated maintenance of chromosome ends. The telomeres are composed of tandem head to tail arrays of repeated sequences made up of two non-long terminal repeat (non-LTR) retrotransposons which orient with the poly-A sequence toward the centromere. The two elements that make up the telomeres are HeT-A (approximately 6 kb) and TART (approximately 10 kb). The retroelements are transcribed into a RNA that is reverse transcribed onto the chromosome ends in much the same way as telomerase adds simple sequence repeats. The elements utilize a reverse transcriptase that uses the 3' end of the nicked DNA as a primer to copy the element into the target site (BIESSMANN *et al.* 1992; LEVIS *et al.* 1993).

Proximal to the telomere is the subtelomere. This region is rapidly evolving and composed of highly polymorphic repetitive DNA (BROUN *et al.* 1992; EICHLER and SANKOFF 2003; MEFFORD and TRASK 2002). The subtelomere is commonly divided into the distal and proximal domains. The distal domain, immediately proximal to the telomere repeats, is often devoid of genes and the repetitive sequences in this domain are commonly found at more than one subtelomere (AMARGER *et al.* 1998; PEARCE *et al.* 1996). The proximal domain is frequently composed of genes or repeats which are found at a smaller subset of chromosome ends (LOUIS 1995; THOMPSON *et al.* 1997). Subtelomeres often contain genes important for adaptation to an organism's environment, such as genes involved in carbon utilization, tolerance to toxic environments, and evading host detection (CHARRON *et al.* 1989;

RACHIDI *et al.* 2000; WADA and NAKAMURA 1996). *M. oryzae* subtelomeres are enriched in genes that encode secondary metabolites and hypothetical secreted proteins that could be involved in pathogenesis (REHMEYER *et al.* 2006). Subtelomeric regions may be unstable due to gene conversion and other types of rearrangements during meiosis (CHUMA *et al.* 2011a). The instability of these regions may allow for rapid evolution of genes involved in pathogenicity, which has been previously observed with instability of Avr-Pita genes (CHUMA *et al.* 2011b).

Chromosome ends in GLS isolates, which are virulent on perennial ryegrass, show "hypervariability" in telomeric RFLPs, while the RB isolates did not show this same pattern. This suggested that there might be structural differences at the chromosome ends between RB isolates and GLS isolates that are leading to the instability. To examine whether there was an underlying structural basis for the telomere variability, homologous chromosome ends were compared between the RB isolate 70-15 and the GLS isolate FH. Figure 1-1 compares four different chromosome ends between FH and 70-15. In general the subterminal regions of the FH isolate are devoid of the transposable elements present in the telomeres of 70-15, with the exception of a RETRO5 long terminal repeat (LTR) element near telomere 8. Telomere 3 in 70-15 is composed of an rDNA array that ends in telomere repeats, while in FH the rDNA array ends in a telomere repeat followed by two copies of a 4.6 kb element and a 2.5 kb element. The two 4.6 kb elements and 2.5 kb element are surrounded by telomere repeats (Figure 1-1, Panel A). In telomere 5, 70-15 contains many transposable elements in the subterminal region that appear to be absent in the corresponding region in FH. The subtelomere of

telomere 5 in 70-15 contains a telomere-linked helicase (TLH) gene. The TLH gene is missing in the corresponding chromosome end in FH, and within the telomere repeat is one copy of a 5.0 kb element in the FH telomere. In telomere 8 of 70-15, the subterminal region contained several transposable elements that were absent in FH. The FH sequence diverged 209 bp from the proximal subtelomere region in 70-15, and there was approximately 2.5 kb of sequence followed by a RETRO5 element. In the telomere of FH there was an element that was 4.2 kb in length followed by three elements that were 1.7 kb in length directed in a head-to-tail orientation towards the chromosome end. In telomere 11 of 70-15, the structure of subterminal region was very different from FH, and contained several different types of transposable elements. The FH subterminal region aligned with disparate regions of the 70-15 genome. This telomere in FH contained three copies of a 5 kb element, the most distal of which, was joined to a truncated 1.7 kb element. (STARNES *et al.* 2012)

The repeated elements embedded in the telomere repeats of the FH isolate were initially defined as *Magnaporthe oryzae* Telomeric Exclusive Repeats (MoTeRs) (FARMAN 2007). MoTER1 was initially described as 4.6 kb in length, and has an open reading frame with a reverse transcriptase domain. MoTER1 lacks long terminal repeats, and was suggested to resemble non-LTR retrotransposons. MoTER2 is only 1.7kb in length, encodes a 204 amino acid hypothetical protein, and lacks a reverse transcriptase domain. They are always found with the 5' end of an element directed toward the telomere. A single MoTeR is embedded in the telomere sequence at some chromosome ends, while in other telomeres tandem arrays of

MoTeRs can be found. Tandem arrays could be made up of single MoTeRs or a mixture of both MoTER1 and MoTER2 (FARMAN 2007). There was limited information available on the distribution of MoTeRs when this work began. It was known that full length MoTeRs were not present in Oryza specialized isolate 70-15. MoTeRs were present in a few Lolium-specialized and Triticum-specialized isolates (Farman personal communication). Additionally, there had been no published studies describing the relatedness of the MoTeRs to the other retrotransposable elements within *M. oryzae*.

The presence of MoTeRs in the telomeres of the GLS isolate and absence in the RB isolate could suggest that they play a role in the instability of telomeres. It is intriguing that the subtelomeres of the GLS isolate FH is devoid of most of the other transposable elements that were found in the RB isolate. One might speculate that the absence of these transposable elements in this region might lead to more stability. Thus it is still unclear what role the different transposable elements may play in the stability of the chromosome ends in the GLS isolates.

1.7 Variability of GLS pathogens and the presence of MoTeRs

MoTeRs were present in GLS isolates, but they were absent in the telomeres of the rice-infecting laboratory strain that has been sequenced. The presence of the MoTeRs at telomeres in GLS isolates is intriguing, as the telomere restriction profiles within GLS populations appear to be highly variable, while the internal chromosome regions show low variability (FARMAN and KIM 2005). This could indicate that the MoTeRs are responsible for the telomere variability observed in the GLS isolates. MoTeRs were not in the genome of the rice-infecting strain 70-15,

based on a preliminary RFLP analysis, and other rice-infecting field isolates did not have high variability at chromosome ends (FARMAN 2007).

The population dynamics underlying these field isolates are not known, nor are the potential differences in adaptive pressures that could affect telomere variability. Rice is a self-fertilizing species that typically grows in a monoculture agricultural environment and rice blast is typically an epidemic disease. These two factors could cause frequent bottlenecking of the rice blast populations leading to low variability. Gray leaf spot is an endemic disease in turfgrasses. Additionally, perennial ryegrass is an obligate outcrossing species (more recombination leading to higher genetic diversity) and is not typically grown as a single genotype in monoculture (higher genetic diversity). Thus GLS pathogens are likely faced with more genetic variability in their hosts, which could possibly lead to higher genetic variability in the pathogen populations. This suggests that the observed telomere restriction profile variability could be based simply on the host's population dynamics. Host diversity has been linked to increased genetic diversity in *M. oryzae* because GLS isolates on tall fescue (another obligately outcrossing species) had more haplotype diversity than St. Augustinegrass (typically vegetatively propagated) specialized isolates (TREDWAY *et al.* 2005). While there could be more variability in the GLS pathogens of perennial ryegrass based just on population dynamics, a mechanism for increased telomere instability cannot be ruled out.

The extreme telomeric variation in the GLS isolates and the presence of MoTeRs suggested that these elements could provide a mechanism by which variability in *M. oryzae* populations might arise in the GLS isolates virulent on

perennial ryegrass. Based on the information outlined above, the focus of the dissertation was narrowed to characterizing some of the mechanisms that could be involved in generating the "hypervariability" in telomere restriction profiles observed in GLS isolates.

1.8 Overview of dissertation research

The overall aim of this dissertation was to characterize the involvement of MoTeRs in telomere instability. Their perceived importance in telomere instability of GLS isolates stems from the facts that: 1) they are located at the telomeres in GLS isolates, 2) transposable elements have been shown to cause instability in *Magnaporthe oryzae,* and 3) the telomeres exhibit extreme instability.

Prior to beginning the experiments outlined in the dissertation (Chapter 2 for general methods and within chapters for specific experimental methods) it was known that there was a MoTeR1, which is described as 4.6 kb in length with an ORF containing an RT domain, and a MoTeR2 that is described as 1.7 kb in length with an smaller ORF of no known function (FARMAN 2007). The MoTeRs do not have long terminal repeats, which suggested that they were a type of non-LTR retrotransposon though the relationship of MoTeRs to other retrotransposons in *M. oryzae* was not known.

The experiments described in Chapter 3 sought to define more clearly the structure of MoTeRs within the GLS isolates, and to determine their relationship with other non-LTR retrotransposons.

While it is interesting to postulate possible mechanisms, such as MoTeR transposition or recombination, that could generate telomere instability in

populations of GLS isolates, I needed to know whether the telomeres were more unstable in GLS isolates than in other host-specialized types of *M. oryzae* without MoTeRs in their telomeres. Previous work had shown that RB isolates of *M. oryzae* from the field showed less telomere variation than GLS isolates (FARMAN 2007; FARMAN and KIM 2005), and sequencing the telomeres of the Oryza isolate 70-15 failed to yield any remnant of MoTeRs. Given this information I designed experiments using Southern hybridization techniques, outlined in Chapter 4, that compared the stability of telomeres in the mitotic progeny of an Oryza pathotype isolate (70-15) to the stability of telomeres in mitotic progeny of a Lolium pathotype isolate (LpKY97-1A) after two infection cycles in their respective hosts. To determine whether instability of GLS telomeres was a general phenomenon among other Lolium specialized isolates, I examined the telomere stability in mitotic progeny from two other GLS isolates.

The Southern hybridization experiments, in Chapter 4, revealed higher telomere instability in GLS isolates. I wanted to determine if at least some of this instability could be due to MoTeRs. To accomplish this, the experiments outlined in Chapter 5 utilized shotgun-cloning followed by Southern hybridization and sequence analysis to follow changes in individual chromosome ends among different mitotic progeny of LpKY97-1A.

Finally, I wanted to determine how widespread MoTeRs were among different host specialized types of *M. oryzae*, as many transposons have an uneven distribution (see above). The only distribution data available for MoTeRs was from a preliminary study that showed that they were not in the rice-infecting laboratory

reference strain (70-15), and were present in a few Lolium and Triticum specialized isolates (Farman personal communication). In Chapter 6, the experiments investigated the distribution of MoTeRs within *M. oryzae* by using Southern hybridizations. The likely evolutionary relationship of MoTeR1 between isolates was determined by sequencing and evolutionary analysis of MoTeR1s from isolates of various host-specialized forms.

Table 1-1. Transposable elements in Magnaporthe.

Class	Subclass	Type	Element name	Distribution within *M. oryzae* pathotypes	Citation
Retrotransposon	LTR	Gypsy	Grasshopper	Present in subgroup of Eleusine isolates	(DOBINSON *et al.* 1993)
			MAGGY*	Absent in Triticum isolates, but present in at least some members of the other host forms	(ETO *et al.* 2001; FARMAN *et al.* 1996b)
			MGLR-3*	Present in all pathotypes tested	(KANG 2001; VIJI *et al.* 2001)
			Inago1	High copy in Digitaria[b], Lolium, Oryza, Setaria Low copy in Eleusine	(SANCHEZ *et al.* 2011)
			Pyret	Present in Eleusine, Eragrostis, Avena, Leersia, Oryza, Pennisetum, and Triticum	(NAKAYASHIKI *et al.* 2001b)
			Retro6	Unknown	
			Retro7	Unknown	
		Copia	Retro5	Present in Digitaria[b], Eleusine, Lolium, Panicum, Paspalum, Panicum, Pennisetum, and Triticum	(FARMAN 2002)
	LINE		MGL(MGR583)	Highest copy number in Oryza and Setaria	(ETO *et al.* 2001)
			MoTeRs	Limited to Lolium and some Triticum isolates[a]	
	SINE		Mg-SINE	Present in most host specific forms with higher copies in Oryza related isolates	(ETO *et al.* 2001)
			MINE	High copy number in Oryza and Pennisetum Single copy in an Eleusine isolate	(FUDAL *et al.* 2005)
Inverted Repeat DNA Transposon			Pot2	Present in most host specific forms	(ETO *et al.* 2001)
			Pot3(MGR586)*	Highest copy number in Oryza	(FARMAN *et al.* 1996a)
			Pot4	Unknown	
			Occan*	High copy number in Oryza Low copy number in Setaria, Eleusine, and Digitaria[b]	(KITO *et al.* 2003)

* Indicates transposable elements implicated in Avr gene instability. Superscript (a) indicates what was known before this dissertation see chapter 5 for updated analyses. Superscript (b) indicates these isolates are from the species *Magnaporthe grisea.*

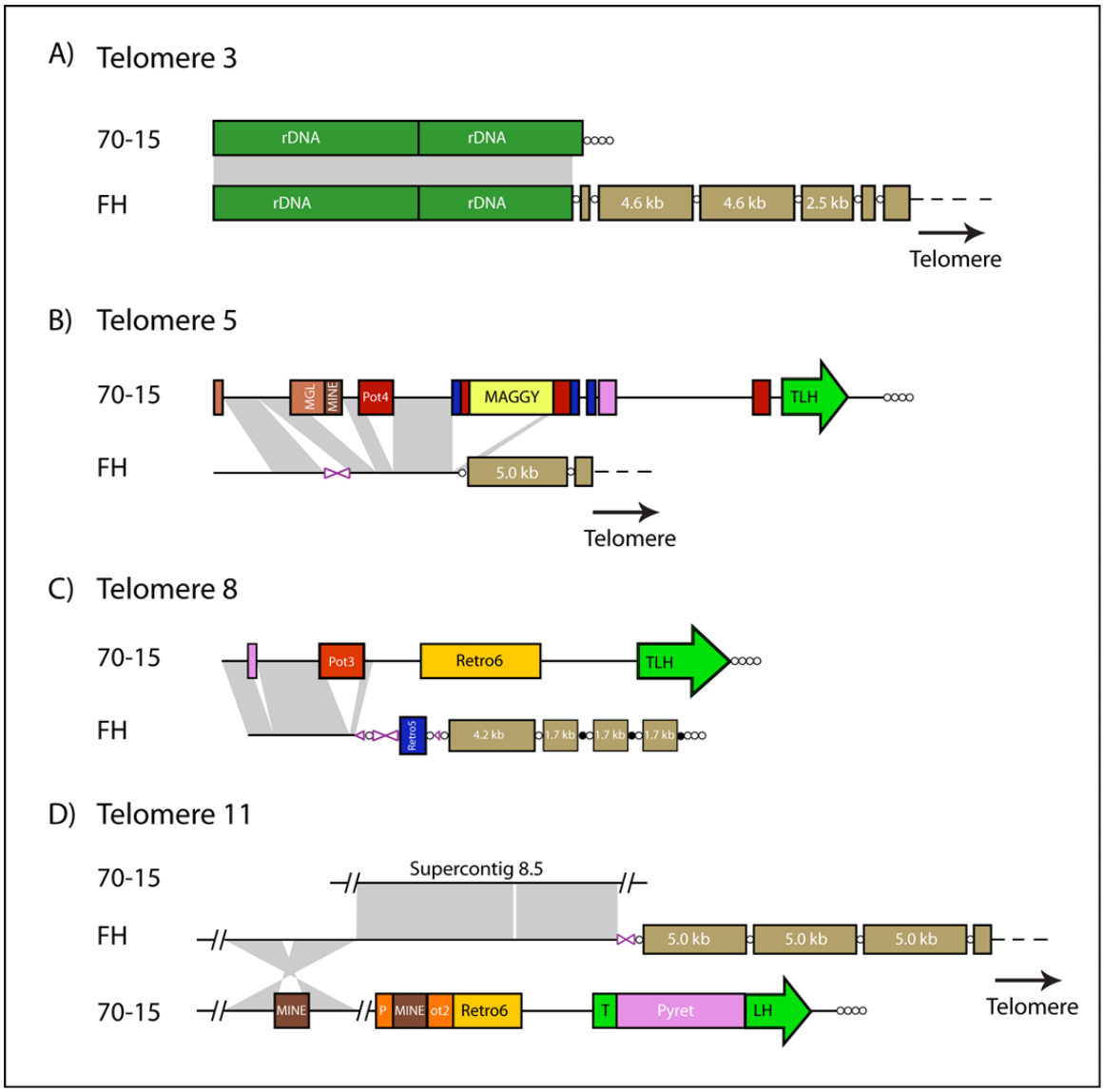

Figure 1-1. Comparison of homologous chromosome end structure in FH and 70-15. Four pairs of chromosome ends are shown in panels A-D. The telomere and the telomere-like sequences are represented with circles, open and black respectively. The direction of the telomere is indicated by a black arrow in FH contigs that did not end in the terminal telomere repeat array. Features discussed in the text are labeled. Repetitive elements in FH telomeres in light brown boxes represent elements that were designated as MoTeRs. Light grey shaded regions indicate alignment of sequence between 70-15 and FH. The FH telomere 11 had alignments with two disparate regions in 70-15, so both of those alignments are shown. Purple triangles represent sequences that were present in FH, but absent in the 70-15 genome. Figure is modified from (STARNES *et al.* 2012).

CHAPTER TWO

Materials and Methods

2.1 Fungal cultures

Fungal strains of *Magnaporthe oryzae* or *Magnaporthe grisea* used in the study are listed in Table 2-1. Fungal isolates were reactivated on oatmeal agar (OA), and grown under constant fluorescent illumination for 7 d. Conidia were harvested by adding 10 ml of 0.1% gelatin and gently rubbing the plate with a bent glass rod. The conidial suspensions were filtered through two layers of cheesecloth, adjusted to 1 x 10^5 conidia per ml by using a hemocytometer, and then 200 µl of the spore suspension was spread on a 4% water agar (WA) plate. This plate was then placed at room temperature overnight in the dark. Germinated single spores were visualized under a dissecting microscope, cut out using a scalpel blade, and transferred to a fresh OA plate. After approximately 3 d, an OA agar plug containing newly grown mycelium was removed from an area adjacent to the single spore. The plug was placed in a test tube containing 10 ml of complete media (CM) for 7-10 days under continuous shaking at room temperature to allow for growth of the culture.

2.2 Plant inoculation and conidia collection

Single spore cultures of 70-15 and LpKY97-1A were grown on OA plates for 7 d. Conidia were collected by flooding the OA plate with 10 ml of 0.1% gelatin and gently rubbing the plate with a bent glass rod. The suspensions were filtered through two layers of cheesecloth. Conidia were counted using a hemocytometer and the conidial suspensions were adjusted to a final concentration of 1 x 10^5

conidia per ml with 0.1% gelatin. Conidial suspensions were sprayed on plants using an artisan's airbrush, after they had been placed into separate plastic inoculation bags. The plants in their respective sealed inoculation bags were placed at room temperature in the dark for approximately 18 h, and then placed in a growth chamber with a 12 h day/night cycle at 27°C day and 21°C night (modified from (KERSHAW *et al.* 1998)). The bags were partially opened for a period of approximately four hours to allow the humidity to equilibrate. Plants were then removed from the bags and infections were allowed to proceed for seven days. Leaves showing lesions were clipped and placed in a moist chamber to sporulate. After 3-5 d in the moist chamber, spores were removed from the lesions using a dry glass rod, and then suspended into 2 ml of a 0.1% gelatin solution.

2.3 Small scale DNA extraction

Mycelium was grown at room temperature with shaking for 7-10 d. The mycelial ball removed from the 10 ml CM using sterilized forceps, and blot dried on paper towels. The dried mycelium was placed in a 96 well round bottom natural polypropylene Whatman uniplate (Florham Park, NJ), snap frozen in liquid nitrogen, and placed in a freeze dryer for 24 h. Mycelium was ground to a powder by shaking using steel beads in a 2000 GenoGrinder (SPEX Certiprep, Metuchen, NJ). Lysis buffer (0.5 M NaCl, 1% sodium dodecyl sulfate [SDS], 10 mM Tris-Cl, pH 7.5, 10 mM EDTA) was warmed to 65°C. One milliliter of preheated lysis buffer was added to the ground mycelium and incubated for 30 min at 65°C. Then 2/3 volumes of (24:24:1) phenol:chloroform:isoamyl alcohol was added, and samples were incubated for another 30 min at 65°C. The plate was then centrifuged for 30 min at 3000 rpm to

pellet the cell debris. After centrifugation, 400 µl of supernatant was transferred to a new 96 well plate, and 240 µl of isopropanol was added. The DNA was then pelleted by centrifugation for 20 min at 3000 rpm. The supernatant was decanted and the pellet was washed twice with 70% EtOH. The DNA pellets were then dried, and redissolved in TE buffer (10 mM Tris-Cl, 1 mM EDTA, pH 8.0) containing RNase A (100 µg/ml). DNA was quantified using SYBR® Gold nucleic acid gel stain (Molecular Probes, Inc., Eugene, OR) in a FluorImager 595 (Molecular Dynamics Inc., Sunnyvale, CA).

2.4 Large scale DNA preparation

In some experiments larger quantities of DNA was needed. Fungal cultures were activated and single spore cultures were generated as described above in section 2.1. These cultures were allowed to grow for three days after which a small agar plug was removed and placed in 50 ml of CM. The culture was shaken for 2 d at room temperature. Then the cultures were blended using a blender, an additional 50 ml of CM was added, and they were placed back on the shaker for an additional 3 d. The fungal cultures were then collected using vacuum filtration, placed in 50 ml BD Falcon™ conical tubes (Becton, Dickinson, and Company, Franklin Lakes, NJ) and snap frozen in liquid nitrogen. The samples were placed in a freeze dryer for 24 h. Mycelium was ground to a powder using a glass rod. The DNA was extracted using the procedures discussed above, but scaled up 10 fold.

2.5 End-enriched clone library preparation

Differential ethanol precipitation was used to remove polysaccharides from the large scale genomic DNA preparations (AMASINO 1986). After purification, the DNA

was redissolved in TE buffer at a concentration of 1 μg/μl. The quality of DNA recovery was checked on a 0.7% agarose gel run at 40 v for 160 min in 0.5X Tris-borate EDTA buffer (TBE [44.5 mM Tris-borate and 1 mM EDTA]). The genomic DNA samples (~ 2 μg) were end-repaired to produce blunt ends using the End-it Repair Kit (Epicenter Technologies, Madison, WI). The end-repaired DNA was then ligated to *Eco*RV-digested pBluescript KS II$^+$ vector (~50 ng) in a 10 μl volume. The ligase was heat-inactivated by incubation at 70°C for 20 min. The samples were digested overnight at 37°C with 20 units of *Pst*I following the manufacturer's protocol (New England Biolabs, Beverly, MA). The *Pst*I was heat-inactivated by incubation at 70°C for 20 min. The reaction mix was diluted 10-fold with 1X ligation buffer, and the *Pst*I digested DNA was then re-ligated using T4 DNA ligase (New England Biolabs, Beverly, MA). Ligase was heat-inactivated by heat treatment at 70°C for 20 min. After ligation, excess salts were removed by dialyzing against TE buffer at 4°C for 1 h. Then the end-enriched plasmid preparation was transformed into the *E. coli* strain EPI300 by electroporation. Cells were incubated for 1 h in SOC media, and 100 μl of the samples were spread onto LB plates with selection for ampicillin resistance (added to the final concentration of 100 μg/ml).

Colonies that contained telomeres were identified by colony hybridization using a $(TTAGGG)_{200\text{-}300}$ probe (FARMAN 2011). Colonies that were positive for telomeres via Southern hybridization were picked and transferred to a 10 ml tube containing LB and shaken overnight at 37°C. Plasmids were extracted using the Zyppy Plasmid Miniprep Kit II (Zymo Research, Orange, California).

2.6 Cloning of MoTeR Bands unlinked to telomeres

A pBluescript KS II$^+$ plasmid vector (Stratagene, La Jolla, California) was linearized using *Pst*I. The sample was run at 80 v for 80 min on a 0.7% agarose gel in 0.5X TBE using 1 kb Plus DNA Ladder^{Tm} (Life Technologies, Carlsbad, CA) as a molecular size marker. Linearized plasmid was excised from the gel, and purified using Qiaquick columns (Qiagen, Valencia, CA). Genomic DNA was digested with *Pst*I, and electrophoresed in 0.7% agarose at 35 v for 24 h in 0.5X TBE. The area of the agarose gel corresponding to the approximate size of the non-telomeric MoTeR was extracted, and purified using Qiaquick columns (Qiagen, Valencia, CA). The linearized plasmid was ligated to the genomic DNA preparation using T4 DNA ligase. In order to remove excessive salts, samples were transferred to a dialysis column, and incubated at 4°C for 1 h in TE buffer pH 7.5. Then 1 µl of the ligation mix was transformed into 50 µl of Transformax EPI300 electrocompetent cells (Epicenter, Madison, WI) by electroporation. Subsequently 950 µl of SOC was immediately added, and cells were incubated for 1 h at 37°C. One hundred microliters of the sample were spread onto LB plates containing ampicillin (100 µg/ml) and X-gal (5-bromo-4-chloro-3-indolyl-β-D- galactopyranoside at a concentration of 40 µg/ml). Colony lifts were performed using Whatman filter paper. Colonies that were positive for telomere sequence via Southern hybridization were picked and transferred to a 10ml tube containing LB and shaken overnight at 37°C. Plasmids were extracted using the Zyppy Plasmid Miniprep Kit II (Zymo Research, Orange, CA).

2.7 Primer design

Primers were designed using VectorNTI version 7 (Invitrogen Corporation, Carlsbad, CA). Primers are listed in Table 2-1. Primers were ordered from Integrated DNA Technologies (Coralville, IA).

2.8 Hybridization probes

The telomere probe was generated by PCR using two primers (TelomereF and TelomereR) with no template. The parameters used in the PCR cycling were: 94°C for 5 min, followed by 35 cycles of 94 °C for 30 s, 50°C for 30 s, and 72°C for 1 min. The final extension phase was at 72°C for 2 min. The reaction products were separated by electrophoresis on a 0.7% agarose gel in 0.5xTBE at 60 v for 120 min. Fragments ranging from 1.5-2.0 kb were excised from the gel, and were then gel purified using Qiaquick columns (Qiagen, Valencia, CA).

MoTeR probes were generated as follows: The primers MoTeR1003F and MoTeR1003R were used to generate the MoTeR1 probe by PCR using a highly diluted plasmid clone as a template. The MoTeR1(RT) probe was amplified by PCR from highly diluted plasmid clones as a template with the primers MoTeR1001F and MoTeR1001R. MoTeR2 template was amplified by PCR from highly diluted plasmid clones as a template with the primers MoTeR2002F and MoTeR2002R. The 5'MoTeR probe was amplified by PCR using the primers MoTeR2001F and MoTeR2001R. The parameters used in PCR cycling were: 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 56°C for 60 s, and 72°C for 1 min, with a final extension phase of 72°C for 5 min. The PCR products were fractionated on 0.7% agarose gel in 0.5X TBE at 80V for 80 min. After electrophoresis, the agarose gel was stained for

30 min in EtBr and 0.5X TBE. Staining solution was removed, fresh 0.5X TBE was added, and the gel was destained for 20 min. The bands corresponding to MoTeRs were excised and extracted using the Qiaquick gel extraction kits (Qiagen, Valencia, CA).

Genomic Probes were generated as follows: Probe 117B was amplified by PCR from the highly diluted plasmid clone 117B1 with the primers m117B1F and m117B1R using ExTaq polymerase. Probe 31B001 template was amplified from plasmid 31B by PCR with the primers 31B001F and 31B001R using ExTaq polymerase. Probe 31B002 was amplified by PCR from plasmid 31B using ExTaq polymerase and the primers 31B002F and 31B002R. The parameters used in PCR cycling were: 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 60°C for 60 s, and 72°C for 1 min, with a final extension phase of 72°C for 5 min. The PCR products were fractionated on 0.7% agarose gel in 0.5X TBE at 80V for 80 min. After electrophoresis, the agarose gel was stained for 30 min in EtBr and 0.5X TBE. Staining solution was drained, fresh 0.5X TBE was added, and the gel was destained for 20 min. The band corresponding to the genomic probe was excised and extracted using the Qiaquick gel extraction kits (Qiagen, Valencia, CA).

Purified amplification products were labeled with Redivue α-32P dCTP (GE Healthcare, Buckinghamshire, United Kingdom) using the Prime-a-Gene Labeling System (Promega, Madison, WI).

2.9 Southern hybridization

One microgram of genomic DNA in a total reaction volume of 50 µl was digested with 20 units of restriction enzyme in the appropriate reaction buffer supplied by

the manufacturer (New England Biolabs, Inc., Ipswich, MA). Approximately 400 ng of digested genomic DNA was loaded onto a 0.7% agarose gel, and DNA was separated electrophoretically at 35 V in 0.5x TBE for 24 h or 48 h at 4°C. The agarose gel was then stained in 0.5% TBE for 30 min in EtBr added to a final concentration of 5 μg/ml, and destained in fresh 0.5X TBE for 20 min. The gel was electroblotted onto a Pall Biodyne B charged nylon membrane (PALL Life Sciences, Ann Arbor, MI) using a GENIE electroblotter (Idea Scientific, Minneapolis, MN) for 2 h at 12 V. The membrane was placed into a small container and the membrane bound DNA was denatured for 10 min in 0.4N NaOH. The 0.4N NaOH was drained, and the membrane was neutralized for 10 min in 2X SSC (1X SSC [saline-sodium citrate] is 0.15 M sodium chloride and 0.015 M sodium citrate). DNA was then UV fixed to the membrane by using the optimal crosslink setting in a Spectrolinker (Spectronics Corporation, Westbury, New York). The membrane was prehybridized for 30 min in hybridization buffer (0.125 M $NaHPO_4$, pH 6.2, 7% SDS, and 1 mM EDTA) at 65°C. The probe was denatured with 0.1 vol. 2N NaOH for 8 min, and then it was neutralized with 0.1 vol. 1 M Tris-Cl pH 7.4. The hybridization buffer used in prehybridization was decanted and replaced with 5 ml of fresh hybridization buffer. Then probe was added and incubated at 65°C for 24 h. Blots were washed twice with 2xSSC for 30 min at 65°C, and then washed once with a high-stringency wash (0.1% SSC and 0.1% SDS) for 30 min at 65°C. The membranes were blotted dry, wrapped in Saran^{Tm} wrap (S. C. Johnson & Son Inc., Racine, WI), exposed on Storage Phosphor Screens (Molecular Dynamics, Sunnyvale, California) at room temperature

for 3 d. The screens were scanned using a Typhoon PhosphorImager (GE Healthcare, Buckinghamshire, United Kingdom).

2.10 Cloning of MoTeR-to-MoTeR junctions

The junctions between MoTeRs were amplified by PCR using different combinations of the forward and reverse primers: MoTeRJ1F, MoTeRJ2F, MoTeRJ1R, and MoTeRJ2R. ExTaq polymerase was used. ExTaq creates a mixture of blunt ended PCR amplicons and 3′ A overhangs in PCR amplicons which allows a PCR amplicons to be directly cloned into T-vectors at an approximate 80% efficiency (Takara, Madison, WI). The parameters used for PCR were: 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 60°C for 60 s, and 72°C for 1 min, with a final extension phase of 72°C for 5 min. PCR amplicons were ligated into pGEM T-easy vectors system using the manufacturer's protocols (Promega, Madison, WI). Transformax EPI300 electrocompetent cells (Epicenter, Madison, WI) were electroporated with plasmid, incubated for 1 h in SOC media, and plated onto LB plates with selection for ampicillin resistance (100 μg/ml) and blue/white colony screening using X-gal (40 μg/ml). Plates were incubated at 37°C overnight. White colonies were picked and transferred to LB broth + amp and incubated overnight. The plasmids were then extracted using the Zyppy Plasmid Miniprep II kit (Zymo Research, Orange, California) or by using Qiagen plasmid preparation kits (Valencia, California) with a Whatman 96 Well lysate clarification UNIFILTER (Florham Park, New Jersey).

2.11 Sequencing

Purified products were sequenced following the protocols outlined in the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA).

Completed BigDye terminated reactions were sent to the University of Kentucky Advanced Genetic Technologies Center (AGTC) for sequencing. Sequences used in subsequent analyses were edited using ContigExpress® in VectorNTI version 7 (Invitrogen, Carlsbad, CA), and derived from forward and reverse sequencing reactions.

2.12 Tandem repeat analyses

Sequences of MoTeR1 and MoTeR2 were analyzed by the program Tandem Repeats Finder version 4 (BENSON 1999). The default parameters were used. Additional editing and analysis was completed manually using VectorNTI version 7 (Invitrogen, Carlsbad, CA). Alignments of MoTeRs were prepared using Kalign (LASSMANN and SONNHAMMER 2005), and manually edited in Jalview version 2 (WATERHOUSE *et al.* 2009) or VectorNTI version 7 (Invitrogen, Carlsbad, CA).

2.13 Relatedness of MoTER1 to other retrotransposons

Full-length protein sequences with reverse transcriptase domains were downloaded from NCBI, or translated from DNA sequence using VectorNTI version 7 (Invitrogen, Carlsbad, CA) where protein sequence was unavailable in the GenBank database (Table 2-3). Putative RT domains were identified using protein BLAST domain search, the protein sequences were then trimmed to remove regions flanking the RT domain and aligned using Kalign (LASSMANN and SONNHAMMER 2005). Where necessary, the multiple alignments were manually edited in Jalview version 2.4 (WATERHOUSE *et al.* 2009), and AlignX® in VectorNTI version 7(Invitrogen, Carlsbad, CA). Phylogenetic analyses using the neighbor-joining method (NJ) was performed using MEGA version 4 (TAMURA *et al.* 2007). Bootstrapping (1,000 replications) was

used to evaluate the statistical support for the NJ tree. The Poisson correction method (PCM) was used to calculate evolutionary distance. PCM calculates distance as $\hat{d} = -\ln(1-p)$ where p is the proportion of sites that differ between two sequences.

2.14 Statistical methods in RFLP comparisons

Analyses of data using one-way ANOVA, two-way repeated measures ANOVA, least square mean t-test post hoc, and two-sample t-tests were completed using SAS v9.2 (SAS Institute, Cary, N.C.). p-values less than 0.05 were considered statistically significant.

2.15 Phylogenetic analyses of MoTeR1 and gene loci

DNA sequences were trimmed manually, and then aligned using Kalign (LASSMANN and SONNHAMMER 2005). Where necessary, the multiple alignments were manually edited in Jalview version 2.4 (WATERHOUSE *et al.* 2009). Phylogenetic analyses using the NJ method was performed using MEGA version 4 (TAMURA *et al.* 2007). Bootstrapping (1,000 replications) was used for statistical support in the trees generated. The Kimura-2 parameter method was used to calculate evolutionary distance.

Table 2-1. Magnaporthe isolates used in the experiments

Isolate Name	Host	Year	Place of Isolation	Source
BdA8401	*Brachiaria distachya*	1984	Los Banos, Laguna, Philippines	R. Nelson
Bm88324	*Brachiaria mutica*	1988	Cabanatuan, Nueva Ecija, Philippines	R. Nelson
DC88428	*Digitaria ciliaris*	1988	Sto. Tomas, Batangas, Philippines	E. Borromeo
217DC	*Digitaria sanguinalis*	2000	Dantzler Ct. Lexington, KY	M. Farman
CG31V01	*Digitaria sanguinalis*	?	State college, PA	W. Uddin
CG31V02	*Digitaria sanguinalis*	?	State college, PA	W. Uddin
DS1100-11	*Digitaria sanguinalis*	2000	Georgia	?
DSDC2	*Digitaria sanguinalis*	1999	Dantzler Ct. Lexington, KY	M. Farman
GGCG1	*Digitaria sanguinalis*	1999	Griffin Gate Golf Course, Lexington, KY	M. Farman
LIZ1	*Digitaria sanguinalis*	2000	Elizabeth St., Lexington, KY	M. Farman
LIZ2	*Digitaria sanguinalis*	2000	Elizabeth St., Lexington, KY	M. Farman
UKCRAB1	*Digitaria sanguinalis*	2008	University of Kentucky	M. Farman
UKCRAB2	*Digitaria sanguinalis*	2008	University of Kentucky	M. Farman
UKCRAB3	*Digitaria sanguinalis*	2008	University of Kentucky	M. Farman
BF17	*Digitaria sp.*	1990	Burkina Faso	D. Tharreau
CD86	*Digitaria sp.*	?	?	D. Tharreau
NJ3	*Digitaria sp.*	2001	Rhode Island	Noel Jackson
NJ4	*Digitaria sp.*	2001	Rhode Island	Noel Jackson
G22	*Eleusine coracana*	1976	Japan	B. Valent
JP29	*Eleusine coracana*	1991	Japan	D. Tharreau
CD156	*Eleusine indica*	1989	Ivory Coast	D. Tharreau
PH42	*Eleusine indica*	1983	Philippines	J. Notteghem
AR4	*Eragrostis curvula*	?	Japan	A. Ellingboe
G17	*Eragrostis curvula*	1976	Japan	B. Valent
LcA8401	*Leptochloea chinensis*	1984	Los Banos, Laguna, Philippines	E. Borromeo
TFMS	*Lolium arundinaceum*	?	Georgia	?

Table 2.1 (continued)

Isolate Name	Host	Year	Place of Isolation	Source
FAMS97-1	*Lolium arundinaceum*	1997	MS	M. Tomaso-Peterson/L. Trevathan
TFRGA	*Lolium arundinaceum*	1997	Georgia	W. Uddin/L. Burpee
PL1	*Lolium multiflorum*	2003	Pulaski County, KY	P. Vincelli
PL2	*Lolium multiflorum*	2003	Pulaski County, KY	P. Vincelli
PL3	*Lolium multiflorum*	2003	Pulaski County, KY	P. Vincelli
CHRF	*Lolium perenne*	1995	Cherry Hill Research Farm, MD	P. Dernoeden
CHW	*Lolium perenne*	1995	Chartwell golf course, MD	P. Dernoeden
FH	*Lolium perenne*	1996	Fountain Head golf course, MD	P. Dernoeden
GG-13	*Lolium perenne*	1997	Griffon Gate golf course, Lexington, KY	M. Farman/P. Vincelli
GG-9	*Lolium perenne*	1997	Griffon Gate golf course, Lexington, KY	M. Farman
HO	*Lolium perenne*	1996	Hanover golf course, PA	P. Dernoeden
KS320	*Lolium perenne*	1996	Manhattan, KS	N. Tisserat
KS330	*Lolium perenne*	1996	Salina, KS	N. Tisserat
LPKY96-1	*Lolium perenne*	1996	Griffin Gate golf course, Lexington, KY	P. Vincelli
LPKY97-1	*Lolium perenne*	1997	Idle Hour golf course, Lexington, KY	P. Vincelli
LPMS97-1	*Lolium perenne*	1997	MS	M. Thomason-Peterson/L. Trevathan
LPOH1	*Lolium perenne*	1997	OH	J. Rimelspach
LPOH3	*Lolium perenne*	1997	OH	J. Rimelspach
NJ1	*Lolium perenne*	1995	New Jersey	B. Clarke
NJ2	*Lolium perenne*	1995	New Jersey	B. Clarke
RGNJ	*Lolium perenne*	1996	Lawrenceville, NJ	S. Vaiciunas B. Clarke
WF9722	*Lolium perenne*	1997	Bloomington, IL	H. Wilkinson
WF9723	*Lolium perenne*	1997	Bloomington, IL	H. Wilkinson
WF9724	*Lolium perenne*	1997	Bloomington, IL	H. Wilkinson

Table 2.1 (continued)

Isolate Name	Host	Year	Place of Isolation	Source
WK3.1	*Lolium perenne*	1996	Yamaguchi, Japan	A. Tanaka/T. Tani
WK3.2	*Lolium perenne*	1996	Yamaguchi, Japan	A. Tanaka/T. Tani
2539	*Oryza sativa*	?	Laboratory Strain	H. Leung
70-15	*Oryza sativa*	1991	Laboratory Strain	A. Ellingboe
82T11	*Oryza sativa*	1982	Texas	M. Levy
91A58	*Oryza sativa*	1991	AR	M. Levy
91L5-1	*Oryza sativa*	1991	LA	M. Levy
91L5-2	*Oryza sativa*	1991	LA	M. Levy
91T14	*Oryza sativa*	1991	TX	M. Levy
91T32-1	*Oryza sativa*	1991	TX	M. Levy
91T32-2	*Oryza sativa*	1991	TX	M. Levy
CH104-3	*Oryza sativa*	?	China	H. Leung
FR10	*Oryza sativa*	1986	Camargue, Domaine GR. Badon, France	J. Notteghem
FR9	*Oryza sativa*	1986	Camargue, France	J. Notteghem
HN1	*Oryza sativa*	1993	Szarvas, Hungry	C. Ipsits
IRRI92328	*Oryza sativa*	1992	Philippines	R. Nelson
IRRI92345	*Oryza sativa*	1992	Philippines	R. Nelson
IT11	*Oryza sativa*	1989	Castle Merlino, Italy	M. Moletti
ML33	*Oryza sativa*	?	Mali	D. Tharreau
R88111	*Oryza sativa*	1988	Cabanatuan, Nueva Ecija, Philippines	E. Borromeo
SIT4	*Oryza sativa*	?	Sitiung, Indonesia	H. Leung
PR8988	*Panicum repens*	1989	Los Banos, Laguna, Philippines	E. Borromeo
PD88413	*Paspalum distichum*	1988	Los Banos, Laguna, Philippines	E. Borromeo
ML36	*Pennisetum*	1991	Mali, Africa	D. Tharreau
US45	*Pennisetum*	?	USA	D. Tharreau
GF1	*Setaria faberii*	1998	KY-American, Lexington, KY	M. Farman
GF2	*Setaria faberii*	1998	KY-American, Lexington, KY	M. Farman

Table 2.1 (continued)

Isolate Name	Host	Year	Place of Isolation	Source
GF2-1	*Setaria faberii*	2000	Elizabeth Street Park, Lexington, KY	M. Farman
GF2000 3-1	*Setaria faberii*	2000	Elizabeth Street Park, Lexington, KY	M. Farman
GF2000 3-2	*Setaria faberii*	2000	Elizabeth Street Park, Lexington, KY	M. Farman
GF9	*Setaria faberii*	1998	KY-American Water, Lexington, KY	M. Farman
STHLAND1	*Setaria faberii*	1998	Southland Park, Lexington, KY	M. Farman
STHLAND2	*Setaria faberii*	1998	Southand Park, Lexington, KY	M. Farman
YF1	*Setaria glauca*	1998	Arboretum, Lexington, KY	M. Farman
YF3	*Setaria glauca*	1998	Arboretum, Lexington, KY	M. Farman
YFDC	*Setaria glauca*	2000	Dantzler Ct., Lexington, KY	M. Farman
SIKY97-2	*Setaria italica*	1997	Caldwell Co., KY	P. Bachi
ARC1	*Setaria viridis*	1998	Arcadia Park, Lexington, KY	M. Farman
ARC2	*Setaria viridis*	1998	Arcadia Park, Lexington, KY	M. Farman
GRF1	*Setaria viridis*	1998	Spindle Top Farm, Lexington, KY	M. Farman
GRF2	*Setaria viridis*	1998	Spindle Top Farm, Lexington, KY	M. Farman
GRF3	*Setaria viridis*	1998	Spindle Top Farm, Lexington, KY	M. Farman
GRF4	*Setaria viridis*	1998	Spindle Top Farm, Lexington, KY	M. Farman
MCCL2	*Setaria viridis*	1998	McConnell Springs, Lexington, KY	M. Farman
PG1054	*Stenotaphrum secundatum*	?	Parrot, GA	L. Tredway
PG1108	*Stenotaphrum secundatum*	?	Lakeland, GA	L. Tredway
SSFL	*Stenotaphrum secundatum*	?	Disneyworld, FL	M. Farman
STAGMS	*Stenotaphrum secundatum*	?	MS	M. Tomaso-Peterson/L. Trevathan
BR115.12	*Triticum aestivum*	1992	Parana, Brazil	Y. Tosa
BR115.7	*Triticum aestivum*	1992	Parana, Brazil	Y. Tosa
BR117.1F	*Triticum aestivum*	1992	Parana, Brazil	Y. Tosa
BR123.1H	*Triticum aestivum*	1992	Parana, Brazil	Y. Tosa
BR131.2	*Triticum aestivum*	1992	Parana, Brazil	Y. Tosa

Table 2.1 (continued)

Isolate Name	Host	Year	Place of Isolation	Source
BR2	*Triticum aestivum*	1990	Parana, Brazil	Y. Tosa
BR202.1B	*Triticum aestivum*	1992	Parana, Brazil	Y. Tosa
BR213.1B	*Triticum aestivum*	1992	Parana, Brazil	Y. Tosa
BR32	*Triticum aestivum*	1991	Brazil	D. Tharreau
BR39	*Triticum aestivum*	1990	Parana, Brazil	Y. Tosa
BR4	*Triticum aestivum*	1990	Parana, Brazil	Y. Tosa
BR49	*Triticum aestivum*	1990	Mato Grosso do Sul, Brazil	Y. Tosa
BR5	*Triticum aestivum*	1990	Parana, Brazil	Y. Tosa
BR52	*Triticum aestivum*	1990	Mato Grosso do Sul, Brazil	D. Tharreau
BR6	*Triticum aestivum*	1990	Parana, Brazil	Y. Tosa
BR7	*Triticum aestivum*	1990	Parana, Brazil	Y. Tosa
BR80	*Triticum aestivum*	1991	Brazil	D. Tharreau
BR81	*Triticum aestivum*	1991	Brazil	D. Tharreau
RN1	*Zingiber sp.*	1986	?	J. Notteghem

Table 2-2. Primers used in the experiments

Primer Name	Length	Sequence 5' to 3'	Tm	Application	Position in MoTeRs
MoTeR1001F	20	CGTTAAAAGACCAGCACGAC	60	MoTeR1(RT) probe	2036-2055[1]
MoTeR1001R	18	TCCCGGCCAATAGACGAA	56	MoTeR1(RT) probe	2591-2574[1]
MoTeR1002F	19	TAACGCCAATACAACCACC	56	MoTeR1(3') probe	4215-4234[1]
MoTeR1002R	19	AGGGACGTTACCTATTACG	56	MoTeR1(3') probe	4685-4667[1]
MoTeR1003F	20	AAGGTATACGCCAAGGCGAT	60	MoTeR1 probe	3125-3144[1]
MoTeR1003R	25	TAGGGACGTTACCTATTACGCACGA	74	MoTeR1 probe	4686-4662[1]
MoTeR2001F	20	AAACCCGAAGGGTTCCCAAG	62	5'MoTeR probe	55-74[1*,2]
MoTeR2001R	25	ATAGATTGGCAACAAAAAGCTACGC	70	5' MoTeR probe	354-330[1,2]
MoTeR2002F	18	AATATTAACGCCGCCGTC	62	MoTeR2 probe	1030-1047[2]
MoTeR2002R	22	TTGGTTAACCCTGAGGATAGAA	62	MoTeR2 probe/Variation	1485-1464[2]
MoTeR1V01F	22	CGGTTTACCCGTTTCTATTAGT	62	Variation	1163-1184[1]
MoTeR1V01R	20	GGCCGGTAATTTTGTCGTGC	62	Variation	2068-2049[1]
MoTeR2V01F	20	ACCTGCTTTTATCACTTGTT	54	Variation	873-892[1]
MoTeR2V01R	22	TGCTGGACTTTTGGCGTAATAA	62	Variation	1324-1303
MoTeRJ1F	25	GGTAACGTCCCTATTTTTGTCTTTG	70	MoTeR Junctions	4674-4698[1]
MoTeRJ1R	22	AAATGCCCACTAATAGAAACGG	62	MoTeR Junctions	1192-1171[1]
MoTeRJ2F	25	CCAACCCAATATATATTCCGATAGG	70	MoTeR Junctions	1584-1608[2]
MoTeRJ2R	19	CAACCCCCCTGTCCAACAG	62	MoTeR Junctions	156-138[1,2]
MoTeRI_F	18	TACGACAAAACCCTTAGC	52	Inverse PCR	4987-5004[1]|1667-1684[2]
MoTeRI1_R	18	TACCTATTACGCACGAGG	54	Inverse PCR	4677-4660[1]
MoTeRI2_R	17	GTTGGTCCTCTTTCGGG	54	Inverse PCR	1603-1587[2]
TelomereF	18	TTAGGGTTAGGGTTAGGG	54	Telomere probe	N/A
TelomereR	15	CCCTAACCCTAACCC	48	Telomere probe	N/A
m117B1F	20	GCCAGCAGACCATATGAACC	62	117B1 probe	N/A
m117B1R	22	CTCGTCTGGAGTGTGTTTAGGG	68	117B1probe	N/A
31B001F	20	AGACGACCGAAACTGTGCAA	60	31B001 probe	N/A
31B001R	20	CCCGTGCACTCTTCTGAAGT	62	31B001 probe	N/A
31B002F	20	TGGTAAATATACTGGGCGAC	58	31B002 probe	N/A

Table 2-2 (continued)

Primer Name	Length	Sequence 5' to 3'	Tm	Application	Position in MoTeRs
31B002R	20	TGCCTTGTATTTCTACCGAC	58	31B002 probe	N/A
117B1JF	20	GTGCGTAATAGGTAACGTCC	60	Unique Junctions	4664-4684[2]
117B1JR	22	ACAATCTGCAAAGGTGCCGCTC	68	Unique Junctions	N/A
MorDNAtelF	19	CGGTAAACTTTACCCAGGA	56	Unique Junctions	N/A
MPG1F	20	AGAAGGTCGTCTCTTGCTGC	54	Evolution MPG1 loci	N/A
MPG1R	19	TTCACTCAACGCTGATCGC	51	Evolution MPG1 loci	N/A
CH7-BAC7F	24	AAGACACGAGAGCAAAGAAAGAAG	54	Evolution CH7BAC7 loci	N/A
CH7-BAC7R	23	CGATACATTACAGTGCCTACGAA	53	Evolution CH7BAC7 loci	N/A
PN3	25	CCGTTGGTGAACCAGCGGAGGGATC	64	Evolution ITS loci	N/A
PN10	22	TCCGCTTATTGATATGCTTAAG	49	Evolution ITS loci	N/A

1 MoTeR1

2 MoTeR2

* Represents the retrotransposon that the probe was amplified from

Table 2-3. Reverse transcriptases used to determine the class of *M. oryzae* retrotransposons.

Name	Organism	GenBank	Reference
Bilbo	*Drosophila subobscura*	AAB92389.1	(BLESA and MARTINEZ-SEBASTIAN 1997)
CR1	*Gallus gallus*	AAA49027.1	(BURCH *et al.* 1993)
Doc	*Drosophila melanogaster*	CAA35587.1	(O'HARE *et al.* 1991)
F	*Drosophila melanogaster*	AAA28508	(DI NOCERA and CASARI 1987)
Jockey	*Drosophila funebris*	P21329	(MIZROKHI and MAZO 1990)
Juan	*Aedes aegypti*	AAA29354.1	(MOUCHES *et al.* 1992)
Lian	*Aedes aegypti*	AAB65093.1	(TU *et al.* 1998)
MGR583	*Magnaporthe grisea*	AAB71689.1	(MEYN *et al.*), (NISHIMURA *et al.* 2000)
NeSL1(Cb)	*Caenorhabditis briggsae*	AAZ15238.1	(NOVIKOVA and BLINOV 2005)
NeSL1(Ce)	*Caenorhabditis elegans*	T25782	(BRADSHAW 1996)
NLR1	*Chironomus thummi*	AAB26437.2	(BLINOV *et al.* 1993)
Q	*Anopheles gambiae*	AAA53489.1	(BESANSKY *et al.* 1994)
R1(Bm)	*Bombyx mori*	AAC13649.1	(XIONG and EICKBUSH 1988)
R1(Dm)	*Drosophila melanogaster*	P16425.1	(JAKUBCZAK *et al.* 1990)
RT1	*Anopheles gambiae*	AAA29363	(BESANSKY *et al.* 1992)
RT2	*Anopheles gambiae*	AAA29365.1	(BESANSKY *et al.* 1992)
SART	*Bombyx mori*	BAA19776.1	(TAKAHASHI *et al.* 1997)
SR1	*Schistosoma mansoni*	AAC06263.1	(DREW and BRINDLEY 1997)
Tad1	*Neurospora crassa*	AAA21792.1	(CAMBARERI *et al.* 1994)
TAHRE	*Drosophila melanogaster*	CAD65869.1	(ABAD *et al.* 2004b)
TART	*Drosophila melanogaster*	T13173	(SHEEN and LEVIS 1994)
TRAS	*Bombyx mori*	BAA07467.1	(OKAZAKI *et al.* 1995)
Dong	*Bombyx mori*	AAA92147	(XIONG and EICKBUSH 1993)
L1(Cl)	*Canis lupus familiaris*	BAA25253.1	(CHOI *et al.* 1999)
L1(Hs)	*Homo sapiens*	AAC51276.1	(SASSAMAN *et al.* 1997)
L1(Mm)	*Mus musculus*	P11369.2	(LOEB *et al.* 1986)
R2(Bm)	*Bombyx mori*	T18197	(BURKE *et al.* 1987)

Table 2-3 (continued)

Name	Organism	GenBank	Reference
R2(Dm)	*Drosophila mercatorum*	AAB94032.1	(MALIK and EICKBUSH 1999)
R4	*Ascaris lumbricoides*	S60004	(BURKE *et al.* 1995)
Swimmer [SW1]	*Oryzias latipes*	AAD02928.1	(DUVERNELL and TURNER 1998)
TX1	*Xenopus laevis*	P14381.1	(GARRETT *et al.* 1989)
ZEPP	*Chlorella vulgaris*	BAA25763	(HIGASHIYAMA *et al.* 1997)
Zorro1	*Candida albicans SC5314*	XP_710131.1	(JONES *et al.* 2004)
Zorro3	*Candida albicans* SC5314	XP_710326.1	(JONES *et al.* 2004)
Cnl1	*Cryptococcus neoformans*	XP_567259.1	(GOODWIN and POULTER 2001)
CRE1	*Crithidia fasciculata*	AAA75435.2	(GABRIEL *et al.* 1990)
CZAR	*Trypanosoma cruzi*	B41950	(VILLANUEVA *et al.* 1991)
FoNLR9	*Fusarium oxysporium*	N/A	(NOVIKOVA *et al.* 2009)
FsNLR1	*Fusarium solani*	N/A	(COLEMAN *et al.* 2009)
Genie	*Giardia intestinalis*	AAL76330.1	(BURKE *et al.* 2002)
GIlT	*Giardia intestinalis*	ABB04054.1	(PRABHU *et al.* 2007)
MoTER1	*Magnaporthe oryzae*	N/A	(FARMAN 2007)
SLACS	*Trypanosoma brucei*	XP_827644.1	(BERRIMAN *et al.* 2005)
Colg2	mitochondrion *Podospora anserine*	CAA38781.1	(CUMMINGS *et al.* 1985)
CoxI	mitochondrion *Marchantia polymorpha*	AAC09454.1	(ODA *et al.* 1992)
Ec107	*Escherichia coli*	CAA44468.1	(HERZER *et al.* 1992)
ECO157	*Escherichia* coli O157:H7 str. Sakai	NP_052642.1	(MAKINO *et al.* 1998)
Mx162	*Myxococcus xanthus*	P23072.1	(INOUYE *et al.* 1989)
Mx65	*Myxococcus xanthus*	P23071.1	(INOUYE *et al.* 1990)
Pstso	chloroplast *Scenedesmus obliquus*	P19593.1	(KUCK 1989)
RET(Ec)	*Escherichia coli*	CAA78293.1	(LIM 1992)
RT67	*Escherichia coli*	P21325.1	(LAMPSON *et al.* 1989)
RT86	*Escherichia coli*	P23070.1	(LIM and MAAS 1989)
1731	*Drosophila melanogaster*	S00954	(FOURCADE-PERONNET *et al.* 1988)
Copia	*Drosophila melanogaster*	P04146.3	(MOUNT and RUBIN 1985)

Table 2-3 (continued)

Name	Organism	GenBank	Reference
RETRO5	*Magnaporthe oryzae*	N\A	(Dean *et al.* 2005)
Ta1-2	*Arabidopsis thaliana*	CAA37920.1	(Konieczny *et al.* 1991)
TNT1	*Nicotiana tabacum*	P10978.1	(Grandbastien *et al.* 1989)
TY1	*Saccharomyces cerevisiae*	P47100.3	(Zagulski *et al.* 1995)
Ty4	*Saccharomyces cerevisiae*	P47024.2	(Janetzky and Lehle 1992)
Ty5	*Saccharomyces paradoxus*	AAC02631.1	(Zou *et al.* 1995)
Grasshopper [GH]	*Magnaporthe oryzae*	AAA21442.1	(Dobinson *et al.* 1993)
MAGGY	*Magnaporthe oryzae*	AAA33420.1	(Farman *et al.* 1996b)
MGLR3	*Magnaporthe oryzae*	AAK01619.1	(Kang 2001)
Pyret2	*Magnaporthe oryzae*	XP_360100.2	(Dean *et al.* 2005)
RETRO6	*Magnaporthe oryzae*	N\A	(Dean *et al.* 2005)
RETRO7	*Magnaporthe oryzae*	N\A	(Dean *et al.* 2005)
Skippy	*Fusarium oxysporum*	S60179	(Anaya and Roncero 1995)
17.6	*Drosophila melanogaster*	P04323.1	(Saigo *et al.* 1984)
297	*Drosophila melanogaster*	P20825.1	(Inouye *et al.* 1986)
412	*Drosophila melanogaster*	P10394.1	(Yuki *et al.* 1986)
Cinful	*Zea mays*	AAD11615.1	(Sanz-Alferez *et al.* 2003)
Gypsy	*Drosophila melanogaster*	P10401.1	(Marlor *et al.* 1986)
RTSb	*Sorghum bicolor*	AAD27571.1	(Llaca *et al.* 1998)
TY3	*Saccharomyces cerevisiae*	Q7LHG5.2	(Hansen *et al.* 1988)
BSOLV	*Banana streak OL virus*	NP_569150.1	(Harper and Hull 1998)
CSSV	*Cacao swollen shoot virus*	NP_041734.1	(Hagen *et al.* 1993)
CSVMV	*Cassava vein mosaic virus*	NP_056848.1	(de Kochko *et al.* 1998)
FMV	*Figwort mosaic virus*	NP_619548.1	(Richins *et al.* 1987)
PCSV	*Peanut chlorotic streak virus*	NP_042513.1	(Richins 1993)
SVBV	*Strawberry vein banding virus*	NP_043933.1	(Petrzik *et al.* 1998)
BFV	*Bovine foamy virus*	NP_044929.1	(Renshaw and Casey 1994)
FFV	*Feline foamy virus*	NP_056914.1	(Bodem *et al.* 1998)

Table 2-3 (continued)

Name	Organism	GenBank	Reference
SFV	*Macaque simian foamy virus*	YP_001961122.1	(Kupiec *et al.* 1991)
BaEV	*Baboon endogenous virus*	P10272.1	(Kato *et al.* 1987)
BIV	*Bovine immunodeficiency virus*	P19560.2	(Garvey *et al.* 1990)
BLV	*Bovine leukemia virus*	P25059.1	(Coulston *et al.* 1990)
CAEV	*Caprine arthritis encephalitis virus*	P33459.1	(Saltarelli *et al.* 1990)
FIV	*Feline immunodeficiency virus*	P31822.1	(Kiyomasu *et al.* 1991)
HIV1	*Human immunodeficiency virus*	P04588.3	(Alizon *et al.* 1986)
HTLV	*Human T-cell lymphotrophic virus*	P03362.3	(Seiki *et al.* 1983)
JDV	*Jembrana disease virus*	Q82851.1	(Chadwick *et al.* 1995)
MLV	*murine leukemia virus*	P03356.2	(Herr 1984)
RSV	*Rous sarcoma virus - Prague C*	P03354.1	(Schwartz *et al.* 1983)
SRV	*Simian retrovirus 1*	P04025.1	(Power *et al.* 1986)
DHBV	*Duck hepatitis B virus* strain China	P30028.1	(Tong *et al.* 1990)
HBV	*Hepatitis B virus*	NP_647604.2	(Okamoto *et al.* 1986)
WHV	*Woodchuck hepatitis virus* 59	P12899.1	(Cohen *et al.* 1988)
TERT(At)	*Arabidopsis thaliana*	AAD54276.1	(Fitzgerald *et al.* 1999)
TERT(Gl)	*Giardia lamblia ATCC 50803*	XP_001709571.1	(Morrison *et al.* 2007)
TERT(Hs)	*Homo sapiens* (human)	NP_937983.2	(Lingner *et al.* 1997)
TERT(Mo)	*Magnaporthe oryzae*	XP_363691.1	(Dean *et al.* 2005)
TERT(Sp)	*Schizosaccharomyces pombe*	AAC49803.1	(Nakamura *et al.* 1997)
Athena(AvO)	*Adineta vaga*	N/A	(Gladyshev and Arkhipova 2007)
Athena(PrB)	*Philodina roseola*	N/A	(Gladyshev and Arkhipova 2007)
Cercyon	*Schistosoma mansoni*	DAA00890.1	(Arkhipova *et al.* 2003)
Coprina	*Phanerochaete chrysosporium*	AAX11377.1	(Goodwin and Poulter 2005)
Neptune(Dr1)	*Danio rerio*	N/A	(Gladyshev and Arkhipova 2007)
Neptune(Dr2)	*Danio rerio*	N/A	(Gladyshev and Arkhipova 2007)
Penelope	*Drosophila virilis*	AAA92124.2	(Evgenev *et al.* 1997)
Xena	*Takifugu rubripes (Fugu rubripes)*	AAK58879.1	(Nogare *et al.* 2002)

CHAPTER THREE

Characterization and Variation of MoTeR Repeats within GLS Isolates

3.1 Introduction

Non-long terminal repeat (non-LTR) retrotransposons have been identified in all major groups of eukaryotes (ARKHIPOVA and MESELSON 2000; GLADYSHEV and ARKIPOVA 2010). Based on phylogenetic analyses of protein-coding sequences, non-LTR retrotransposons have been suggested to be the oldest type of mobile elements that use a reverse transcriptase to create new copies (MALIK and EICKBUSH 2001; MALIK *et al.* 2000).

Functional non-LTR retrotransposons encode proteins with several different enzymatic activities, which allow them to utilize a relatively simple insertion mechanism (CHABOISSIER *et al.* 2000; FENG *et al.* 1996; LUAN *et al.* 1993). First, a chromosomal target site is cleaved by an endonuclease encoded by the retrotransposon. Depending on the element, the endonuclease may be a restriction-like endonuclease (REL-endo) (MALIK *et al.* 1999) or an apurinic/apyrimidinic (APE) endonuclease (ZINGLER *et al.* 2005a). Next, the 3' end of the cleaved DNA is used as a primer for the reverse transcriptase (encoded by the retrotransposon) to polymerize a cDNA copy of the RNA transcript at the target site. This method is typically called target primed reverse transcription (TPRT). The second strand is then synthesized, though the mechanism by which this occurs is still unclear (HAN 2010).

There is considerable structural variability among non-LTR retrotransposons which could cause variation in the precise steps described above (EICKBUSH and MALIK 2002). The oldest lineages of non-LTR retrotransposons have a REL-endo domain located on the carboxy terminal (C-terminal) side of the RT domain, which is encoded by a single open reading frame (ORF). The more recent lineages of non-LTR retrotransposons lack the REL-endo domains and have a more complex structure. They have acquired a second open reading frame (ORF1) that is located upstream of the RT domain (ORF2). ORF1 is thought to play a role similar to gag proteins of retroviruses (DAWSON *et al.* 1997), it has nucleic acid chaperone activity (MARTIN and BUSHMAN 2001), and RNA binding activity (KOLOSHA and MARTIN 2003). In ORF2, the APE domain is located on the amino terminal side (N-terminal) of the RT (MALIK *et al.* 1999).

Two other regions play an important role in non-LTR retrotransposon function, and these are the 5' and 3' untranslated regions (UTRs). The 5' and 3' UTRs are highly variable between various non-LTR retrotransposons. The 5' UTR contains internal promoter activity (MIZROKHI *et al.* 1988), but this promoter can frequently be replaced (HAAS *et al.* 2001; KHAN *et al.* 2006), giving rise to 5' UTR sequence variability among elements from different species. Some variation at the 5' UTR between elements within the same genome may also be due to frequent truncation as a result of incomplete reverse transcription (AKSOY *et al.* 1990; SZAK *et al.* 2002).

The 3' UTR typically has a specific sequence/structure that is recognized by the reverse transcriptase (ANZAI *et al.* 2005; LUAN and EICKBUSH 1995; OSANAI *et al.*

2004). At the 3' boundaries of non-LTR retrotransposons there are one of three different sequence types. In some non-LTR retrotransposons, a variable length polydeoxyadenosine (polyA) sequence is located at the 3' boundary. Two examples of variable polyA sequence at the 3' boundary are the L1 retrotransposon where the polyA tract varied from 27 to 54 nucleotides (DOMBROSKI *et al.* 1991) and the TRAS1 retrotransposon where the length of the poly(A) tract varied from 44 to 72 nucleotides (OKAZAKI *et al.* 1995). Other non-LTR retrotransposons have short sequence repeats (STRs). For example, the non-LTR retrotransposons Dong (XIONG and EICKBUSH 1993), Q (BESANSKY *et al.* 1994), and I (FAWCETT *et al.* 1986) end in TAA repeats, the CR1 retrotransposon ends in the repeat ATTCTGT (BURCH *et al.* 1993), and the R1 retrotransposons end in GTC or TA repeats (EICKBUSH and EICKBUSH 1995). The last type of 3' end found in non-LTR retrotransposons is an apparent absence of either STRs or polyA sequence (HAN 2010).

The *Magnaporthe oryzae* Telomeric Retrotransposons (MoTeRs) are classified as non-LTR retrotransposons due to the lack of long terminal repeats bordering the ends of the repeats. The larger of the repeats (MoTeR1) is 5 kb and has a reverse transcriptase-like domain. MoTeR2 is a smaller repeat being only 1.7 kb in length, and lacks a reverse transcriptase domain. These repeats are found in telomeres and are bordered by telomere repeats (TTAGGG)n (FARMAN 2007; STARNES *et al.* 2012). There was little else known about the specific structure of these elements. The experiments in this chapter sought to address the structure, genomic distribution, and variability of the MoTeRs within GLS isolates of *M. oryzae*.

3.2 Results

3.2.1 MoTeRs structure and classification

To describe the commonality between MoTeR1 and MoTeR2, the DNA sequences were aligned. This revealed that the elements had extensive similarity at their 5' ends, as there are only two mismatches (one mismatch at position 7 where there is a G/A mismatch and another at position 63 where there is a G/A mismatch) in the first 870 bp of alignment. There is also perfect sequence identity of 77 bp at the 3' ends of the two MoTeRs. The sequences did not share similarity outside of these two regions. There were numerous short tandem repeat (STR) motifs within MoTeRs that were found using Tandem Repeats Finder (Table 3-1). Six of the STRs, all located in the 5' UTR, were shared between the elements. The two STRs closest to the telomere, STR(A) [CCCGAA] and STR(B) [CCCAAA], are very similar in sequence to the telomere repeat (CCCTAA)n found in *M. oryzae*. MoTeR1 had five unique STRs, only one of which (repeat I) was located in the ORF. MoTER2 had an additional five unique STRs, three of which were located within the ORF (Figure 3-1).

In the region where no sequence similarity was found between MoTeR1 and MoTeR2, each element had an ORF encoding different proteins. In MoTeR1 an ORF coding for a 1070 amino acid protein was detected, and in MoTeR2 an ORF coding for a 204 amino acid protein was discovered. These ORFs are shown in Figure 3-1 as solid arrows.

Hypothetical protein sequences were translated from the DNA sequence using VectorNTI version 7.0. Blast searches in GenBank using the non-redundant

protein sequence database were then used to determine if there was similarity to other known elements. The MoTER2 predicted protein was not similar to any known protein. A reverse transcriptase domain (Pfam ID: PF00078) was detected within the MoTER1 predicted protein, and this protein showed significant similarity to protein sequence from CNL1, a telomeric retrotransposon from *Cryptococcus neoformans*, the SLACS retrotransposon in *Trypanosoma brucei* and *Leishmania braziliensis*, and CRE1 and CRE2 retrotransposon from *Crithidia fasciculate* (GABRIEL *et al.* 1990). These retrotransposons are described as being in the CRE-like clade of non-LTR retrotransposons. Additionally, MoTeR1 was similar to two other members of CRE-like retrotransposons, which were recently discovered in *Fusarium oxysporum* (FoNLR9) and *Fusarium verticillioides* (FvNLR4) (NOVIKOVA *et al.* 2009). The similarity of the MoTeR RT domain to other non-LTR retrotransposons suggested that the MoTeRs are non-LTR retrotransposons, and consequently the elements were designated *Magnaporthe oryzae* Telomeric Retrotransposons (MoTeRs). This finding was significant as the MoTeRs are the first CRE-like retrotransposons to be found in *M. oryzae*.

The top Blast match for the MoTeR1 RT was to a hypothetical protein in *Nectria haematococca*. Further analysis of the Nectria protein indicated that it was encoded by a previously unknown non-LTR retrotransposon that was present in telomeric and non-telomeric locations in the *N. haematococca* genome. Full-length elements were only present near telomeric locations (data not shown). This retrotransposon was thus named *Nectria haematococca* Telomeric Retrotransposon

(NhTeR1), and the DNA sequence of the full-length (4.3 kb) element is provided in Appendix A.

To understand the relationship of the MoTeR1 hypothetical reverse transcriptase with known reverse transcriptases within *M. oryzae* and other organisms, a phylogenetic analysis needed to be performed. To accomplish this, full-length protein sequences with reverse transcriptase domains were downloaded from NCBI, or translated from DNA sequence where protein sequence was unavailable in the GenBank database (Table 2-3). The putative RT domains were identified using protein BLAST domain search, the protein sequences were then trimmed to remove regions flanking the putative RT domain, and aligned. Where necessary, the multiple alignments were manually edited to align the conserved motifs within the RT (XIONG and EICKBUSH 1990). Phylogenetic analyses using the neighbor-joining method (NJ) were performed. Bootstrapping (1,000 replications) was used to evaluate the statistical support for the NJ tree. The Poisson correction method (PCM) was used to calculate evolutionary distance.

Based on the phylogenetic analysis, the MoTeR1 reverse transcriptase was did not group with any of the previously discovered reverse transcriptases in *M. oryzae*, marked in bold text in Figure 3-2. MoTeR1 was most similar to NhTeR1, Cnl1, and FoNLR9 (Figure 3-2), which are elements also located in or near telomeres in other fungi. This group was closely associated with the relatively ancient CRE clade, which insert in a sequence-specific manner (EICKBUSH 2002). MoTeR1 was somewhat more distantly related to the Giardia early non-LTR insertion elements

(Genie), which inserts into a repeated sequence located near telomeres (BURKE *et al.* 2002).

The CRE-like retrotransposons typically have N-terminal zinc-finger nucleic acid binding motifs with the consensus sequence $C(X_2)C(X_{12})H(X_{4-5})H$, and a C-terminal REL domain with the consensus sequence $C(X_{1-3})C(X_{7-8})H(X_{3-4})C(X_{9-10})RHD/N(X_{19-33})E(X_{9-21})$R/KPD β-turn D/E (where X may be any amino acid). These motifs are believed to be involved in sequence specific integration (CHRISTENSEN *et al.* 2005; YANG *et al.* 1999). Motivated by the similarity of the MoTeR1 RT domain to these retrotransposons, the MoTeR1 protein sequence was analyzed for these motifs. Two putative zinc-finger binding motifs were discovered within MoTER1 with the sequence $C(X_2)C(X_9)H(X_4)H$ and $C(X_2)C(X_{12})H(X_4)H$. The sequences of the MoTeR1, NhTeR1, FoNLR9, Cnl1, SLACS, CZAR, CRE1, and CRE2 were aligned using Kalign, and manually edited to show the alignment of the conserved motifs of the C-terminal REL domain (Figure 3-3). MoTeR1 shared many of the REL domain conserved motifs with other CRE-like retrotransposons. These conserved motifs in MoTER1 are indicated by bold print while the differences are in plain text in the sequence **C**(X_2)**C**(X_7)**H**(X_3)**C**(X_9)**RHD**(X_{18})**E**(X_{48})**R**A**D** β-turn **D**. Most CRE-like transposons have a short, non-conserved region in the REL domain between the EP and the RPD conserved motifs (Figure 3-3), mostly between 8-15 amino acids (i.e. EP(X_{8-15})RPD). MoTeR1 has a large extension of this non-conserved region (X_{47}) whereas other retrotransposons closely related to MoTeR1 have much shorter non-conserved regions (NhTeR1 and FoNLR9 both have X_7). MoTeR1, NhTeR1, and FoNLR9 did not have the characteristic RPD domain

associated with CRE-like retrotransposons, and instead had an RAD domain (Figure 3-3).

3.2.2 Genomic distribution of MoTeRs among GLS isolates

Full length MoTeRs were found in the telomeres of the Lolium pathotype isolate FH, but not in the genome sequence of 70-15. To determine if MoTeRs were present in telomeric locations in other Lolium-specialized isolates, DNA samples were digested with *Pst*I. When probing genomic DNA digested with *PstI* with telomere probe, the hybridizing fragments will include the telomere sequence to first *Pst*I site in from the chromosome end. The MoTeRs do not have *Pst*I sites in their sequence and so the MoTeRs arrays within that fragment should remain intact within the telomere fragments. Following digestion, the DNA was fractionated by electrophoresis, and then electroblotted on a nylon membrane. The blots were then hybridized sequentially with a MoTeR1-specific probe, MoTeR1-specific probe (RT sequence), a MoTeR1-specific (3') probe, a MoTeR2-specific probe, and the telomere probe. Between each of the hybridizations the blots were stripped to remove the previous signals.

Every GLS isolate analyzed with telomere probe in the Southern hybridization had different telomere fingerprint patterns (Figure 3-4-A). The hybridization patterns for the telomere (Figure 3-4-A), MoTeR1 (Figure 3-4-B), and MoTeR2 (Figure 3-4-E) probes revealed that MoTeR1 and MoTeR2 mainly cohybridized with the telomeric fragments and not to other internal *Pst*I fragments. With regards to MoTeR2 there was only one notable exception in which the MoTeR2 did not cohybridize with telomere (marked with a red asterisk in Figure 3-4-E).

There were five discernible exceptions where fragments hybridized strongly to the MoTeR1 probe, but did not hybridize to the telomere probe (highlighted with asterisks in Figure 3-4-B). These were found in the isolates CHRF, KS330, RGNJ, and TFMS. The non-telomere hybridizing MoTeR1 fragment in RGNJ (marked with a red arrow in lane 12 Figure 3-4-B) was cloned and partially sequenced. This revealed 384 base pairs of unknown sequence followed by a highly truncated MoTeR1 on one end of the plasmid insert, and a MGRL-3 on the other side of the plasmid insert (Figure 3-5). A faint, non-telomeric MoTeR1-hybridizing fragment was observed in 10/14 isolates (Figure 3-4B marked with a red arrow in lane 14). This fragment was cloned from TFMS and sequenced. The resulting TFMS fragment was 1886 bp in length (Figure 3-5). The sequence did not contain any telomere repeats indicating it was likely at an internal locus. Blasting this sequence against GenBank indicated sequence similarity, from position 315-629, to the telomeric clone 72H05 located on telomere 12 in 70-15 from position 8024-8339 (REHMEYER *et al.* 2006). There were only 15 mismatches between the sequences. The local blast search of the TFMS fragment sequence to MoTeR1 indicated there was sequence similarity from positions 629-1716 in the TFMS fragment to positions 4021-5035 in MoTeR1. However, there were a high number of mismatches (217) and gaps (15) in the alignment, which results in the faint intensity of the MoTeR1 fragment in the Southern blots.

The Southern hybridization data were tabulated to provide information on MoTeRs in telomeres of the GLS isolates (Table 3-2). Telomere hybridizing *Pst*I fragments were counted as one telomere for the purposes of this experiment. This

method may underestimate the number of telomeres if some telomeric fragments comigrate during electrophoresis prior to Southern analysis. A total of 136 telomere-hybridizing bands were counted with an average of approximately 11 bands per isolate. Most of these telomere-hybridizing fragments cohybridized with at least one of the MoTeRs (123 out of 136), which suggested that most of the telomeres in GLS isolates have MoTeRs. In 6 out of the 14 isolates analyzed all of the visible telomere-hybridizing fragments cohybridized with MoTeRs. Some of the telomere fragments cohybridized with either MoTeR1 or MoTeR2, while others cohybridized with both MoTeRs. There were only 13 telomere-hybridizing fragments that appeared to lack MoTeRs altogether (Table 3-2).

The method used above will underestimate the number of telomeres, as several telomeric *Pst*I fragments may migrate together during electrophoresis due to their similar size. To gain a better estimate of MoTeR copy numbers in the Lolium-specialized isolates, the number of copies was estimated by visual means, taking the relative hybridization intensities into account. Densitometric scanning of the lanes in the Southern blot was also used to provide an independent assessment of copy numbers. These estimations are provided in Table 3-3. As seen in the table, visual estimation tended to underestimate the number of copies present within intensely hybridizing fragments, and thus the densitometric scanning method was considered more reliable in determining the number of copies present based on the intensity of the hybridizing fragments in the Southern Blots. Based on densitometric scanning, there were, on average, ~14 telomeres per isolate with a range from ~10 to ~17. The average copy number of MoTeR1 was 15.5 with a range from 10 to 27.

The copy number of MoTeR2 was lower with an average of 11.3 per isolate, and a range from 0 to 23. CHRF exhibited the highest MoTeR copy numbers with ~50, while PL3 had the least with ~15 MoTeR1s and no MoTeR2s.

3.2.3 Variability of MoTeRs among Lolium pathotype isolates

There is one *Bam*HI site present within MoTeR1 and MoTeR2 (position 4579-4585 in MoTeR1 and position 1020-1025 in MoTeR2). When the genomic DNA is digested with BamHI some of the fragments should contain the terminal MoTeRs with their respective telomere repeats (Figure 3-6). It was believed that variation of the terminal MoTeRs among different Lolium pathotype isolates could be detected by Southern analysis of *Bam*HI digests with a telomere probe. A preliminary analysis of *Bam*HI-digested DNAs from Lolium-specialized isolates with a telomere probe produced a small number of intense signals, which was believed to be due to the presence of conserved MoTeR sequences at multiple chromosome ends (i.e. the terminal MoTeRs). However, the sizes of the intense signals varied between different GLS isolates, implying that different isolates contain different MoTeR variants (data not shown).

To expand upon this finding Southern hybridization analyses of genomic DNA digested with *Bam*HI from different strains was used to detect variation of MoTeRs in the Lolium pathotype isolates using 4 different probes. Additionally, different cultures of the same fungal isolate were used in KS320, LpKY97-1A, LpOH97-1, and RGNJ to see if variation could exist within mitotic progeny of the same isolate. The four probes used in the analyses include: telomere, MoTeR(5'), MoTeR1(RT), and MoTeR2. The MoTeR(5') probe could hybridize to both MoTeR1

and MoTeR2 due to the similarity of the 5' ends. This probe could potentially hybridize to six different fragments: including those that have terminal MoTeRs and various different configurations of MoTeR-MoTeR junctions (See Figure 3-6). The MoTeR1(RT) probe was limited to hybridization with terminal MoTeR1s or in fragments within MoTeR1-MoTeR junctions. The MoTeR2 probe was limited to hybridization with subterminal elements.

When probed with the telomere probe two intense signals were present in 18 of the 27 of the isolates (Figure 3-7-A in the region marked with black boxes). These intense telomere signals cohybridized with the MoTeR(5') probe (Figure 3-7-B). The higher intensity of the signals and the cohybridization of the MoTeR(5') probe indicates that multiple chromosome ends contain the same size terminal MoTeR copies within an isolate. The larger sized fragments (black box MoTeR1 Figure 3-7-A) corresponds with MoTeR1 based on cohybridization of the MoTeR1(RT) (Figure 3-7-B) and MoTeR(5') (Figure 3-7-C) with this fragment, while the smaller fragments (black box MoTeR2 Figure 3-7-A) corresponds with the expected size of the terminal MoTeR2 and cohybridized with MoTeR(5') (Figure 3-7-B). Size variation in the terminal MoTeR copies was observed between isolates (Figure 3-7-A). For example, the terminal MoTeR1 copies in FH (lane 5 Figure 3-7-A marked with a red arrow) are smaller than terminal MoTeR1s in CHW (lane 3 Figure 3-7-A marked with a red arrow). The difference in size variation was also the same for MoTeR2 in these isolates, with the terminal MoTeR2s in FH being smaller than the corresponding band in CHW (lanes 5 and 3 respectively, in Figure 3-7-A marked with a green arrow). The terminal MoTeRs in RGNJ were intermediate in

size between FH and CHW (lane 7 Figure 3-7 A). Since it had already been established that full length MoTeR1 and MoTeR2 share the same 5' ends in the isolate FH (Figure 3-1), the likely explanation of the size variation of terminal MoTeRs among different isolates is due to differences in the shared 5' sequence between MoTeR1 and MoTeR2.

Five isolates (CHRF, FaKY97-1A, KS320, and RCHMD) lacked the intense signal expected from full-length telomeric MoTeR1s, but instead showed multiple bands with weaker signals around the expected size, indicating that these isolates have variation in the terminal MoTeR1s (Figure 3-7-A). Based on the lack of MoTeR(5') hybridization, LpMS97-1 (lane 16) and the WK isolates (lanes 32-34) did not have any full-length MoTeRs in their genomes (Figure 3-7-B). The WK isolates did retain 5' truncated MoTeRs as was seen with hybridization signals to MoTeR1(RT) and MoTeR2 probes (lanes 32-34 Figure 3-7 C and D respectively). The truncated MoTeRs in the WK isolates cohybridized with the telomere probe indicating that truncated MoTeRs are still located at telomeres. In the LpMS97-1 isolate, no hybridization to the MoTeR2 probe was observed, which suggests that it either does not have MoTeR2 or any MoTeR2s are truncated. There was a faint signal with the MoTeR1(RT) probe in LpKYMS97-1A that cohybridized with the telomeric fragment signifying that it did have a truncated MoTeR1. There were no MoTeR hybridization signals in the rice infecting isolate 70-15.

Different laboratory cultures of the same isolate generally showed the same telomere profiles, as seen in strains KS320, LpOH97-1, and RGNJ. LpKY97-1A showed an appearance of a new telomere band in one subculture (marked with a

red asterisk in Figure 3-7-A). This new telomeric fragment cohybridized with MoTeR1(RT), but didn't cohybridize with MoTeR(5') probe. There were several additional new bands in the LpKY97-1A culture that were seen with the MoTeR1(RT) and MoTeR2 hybridizations. While no new telomeric fragments were seen in RGNJ, the MoTeR2 probe revealed some polymorphisms between the cultures (lanes 7 and 8 Figure 3-7-D).

3.2.4 Structural Variability of MoTeRs within FH

Variability in the terminal MoTER1s and MoTeR2s were observed in the Southern Blots of *Bam*HI digested genomic DNA from various GLS isolates (Figure 3-7). To understand the cause of this variability, I first analyzed three size variants of MoTeR1s (4.6 kb, 4.9 kb, and 5.0 kb) that had been captured in telomeric clones from FH and sequenced (sequence data generated by Dr. Farman). The sequences were aligned using VectorNTI and manually edited. These MoTeR1s had variation near the 5' end of the sequence (Figure 3-8). The 4.6 kb MoTeR1 was shorter due to a truncation at the 5' end. The 4.9 kb MoTeR1 was missing one copy of STR(G) and one copy of STR(H) that were found in the 4.6 kb and 5.0 kb MoTeR1s. In addition, the 4.9 kb MoTeR1 had a tandem duplication of the sequence TTACCTGCTTT, which had not previously been identified as a tandem repeat. The 4.6 kb and 5.0 kb element only have one copy of the TTACCTGCTTT sequence. Other than the difference at the 5' end, there are very few mismatches within the different size variants (data not shown).

3.2.5 Factors involved in size variation of MoTeRs

If a GLS isolate had a longer terminal MoTeR1 it also tended to have a longer terminal MoTeR2 indicating that the differences among different isolates could be due to variation in the common sequences at the 5' ends of the MoTeRs. For example, CHW showed longer MoTeR1 and MoTeR2 fragments than LpKY97-1A, which in turn had longer fragments than did FH and RGNJ (Figure 3-7). Variation in short tandem repeats found near the 5' end of the MoTeR1s had already been seen within FH (Figure 3-8), so it was believed that extension or deletion of repeats within the MoTeRs could account for this variation. Extension of any of the 5' end STRs could lead to the size variation observed in Southern analyses of *Bam*HI digests with a telomere probe.

To examine whether variation in 5' STRs could account for size variation MoTeR sequences obtained previously or generated from MoTeR-MoTeR junctions were aligned and analyzed for insertion or deletion of sequences. The new sequence was generated by PCR amplification of MoTeR-MoTeR junctions, which could potentially yield a large number MoTeRs 5' ends with differences in the copy numbers of the 5' STRs. Although the PCR amplicons would not reflect the terminal MoTeRs, the data from the Southern analysis of *Bam*HI digests showed that if an isolate had a longer terminal MoTeR fragment, then this difference was also seen in the comparison between internal MoTeRs. This suggested that the majority of copies in a genome would share the same organization. Forward primers were designed to anneal the 5' end of a MoTeR, and reverse primers were designed to read from the 3' end of another MoTeR.

A total of 370 MoTeR-MoTeR junction plasmid clones were sequenced and analyzed. Thirty-one different junction sequences were obtained and are listed in Table 3-4. At the extreme 5' end of the MoTeRs, telomere repeats are followed by the variant telomere repeats STR(A) and STR(B). There was no consistent difference among isolates in the number of telomeric repeats (CCCTAA) at the MoTeR-MoTeR junctions (Table 3-4). In 28 of the 31 junctions there was at least one full telomere repeat unit between the 3' ends of one MoTeR to the 5' end of the neighboring element. The largest number of telomere repeat units between MoTeRs was 28 (in PL1 and RGNJ). In the three junctions that were missing a telomere repeat adjacent to the 3' end of the MoTeR there were variant telomere sequences (CCCAAA, CCCGAA, or CCCTA).

STR(A), CCCGAA, characterized in MoTeR1 and MoTeR2 in the isolate FH was not found at the extreme 5' ends in the LpKY97-1A isolate. 5' ends captured from LpKY97-1A started with the telomere repeats followed by the variant repeat STR(B), CCCAAA. LpKY97-1A had the greatest number of the STR(B) at the 5' end (11 copies), while PL1 consistently had the lowest number of copies of STR(B). PL1 has a smaller terminal MoTeR1 and MoTeR2 than LpKY97-1A (Figure 3-7). Taking into account the highest copy number of STR(B) in both LpKY97-1A (11) and PL1 (6) the difference based on STR(B) in non-terminal copies is only 30 bp. While there is some variation in at least STR(B), it is not great enough to account for the size difference of MoTeRs as seen in the Southern analyses.

Other 5' end STRs are located downstream of the variant telomere repeats (STR(A) and STR(B)) (Figure 3-1), and are variable both within and among different

isolates (Figure 3-9). There was an extension of the STR(C) in CHW in both MoTeR1 and MoTeR2 sequences (shown in light blue in Figure 3-9-C and 3-9-B respectively), which could account for the larger size seen in the Southern blots. Sequence obtained from LpKY97-1A showed at least four different MoTeR1 elements in various stages of truncation, and are specified by a number in parenthesis beside the name (Figure 3-9-C). In addition a new STR, AAAAGTCGCTAGACTTTTACTAT, which had not been previously identified with Tandem Repeats Finder, was found in the MoTeR1 of plasmid clone LpKY97-1A (2). This STR was repeated 2.8 times in that copy. In the other LpKY97-1A MoTeR1 clones (1 and 3) this STR was repeated 1.8 times. This STR was also extended in the MoTeR2 of CHW and FH. Other smaller repeat extensions were observed. In the MoTeR2 of LpKY97-1A an extension of a poly(T) tract was seen. In FH, RGNJ, KS331, and CHW, two copies of the STR (TTACCTGCTTT) was observed whereas in LpKY97-1A it was not a repeat.

The junction sequences also provided information on the 3' end of the insertions. Some non-LTR retrotransposons will create extended poly(A) tails at the 3' ends post-transcriptionally. The extended repeats are then incorporated upon insertion of the element into the genome, which leads to variable length poly(A) runs at the 3' insertion. This could represent a source of size variation. However, none of the MoTeR junction sequences had extensions or variations in the four adenine bases at the 3' end (5' AATAAAGCGCGAATT**AAAA** 3') indicating that the MoTeRs do not have a poly(A) tail.

While there was variation in the repeats at the 5' end of the element, some size difference could also have been due to a loss/gain or a change in the position of

a *Bam*HI restriction site. Potential loss of the site was tested by PCR followed with restriction digests of the PCR products as outlined below. Sequence spanning the *Bam*HI site in MoTeR1 was PCR-amplified from various GLS isolates using the primers MoTeR1002F and MoTeR1002R. MoTeR2 sequence was amplified using the primers MoTeR2V01F and MoTeR2002R. Based on the MoTeR1 and MoTeR2 sequence of FH (Appendix 1) there should be one BamHI site in each of the PCR amplicons. A total of 8 strains that showed size variation in telomeric MoTeR fragments were used to amplified MoTeR1 and MoTeR2 from genomic DNA using the primer sets outlined above. The PCR amplicons were then digested with BamHI to survey for the presence of the restriction site. The PCR amplicons corresponding to MoTeR1 and MoTeR2 as well as the digested products were separated by electrophoresis on an agarose gel.

The MoTeR1 amplicons were roughly the same size between the different strains (Figure 3-9-A, lanes 2-9), which indicates that size variation in the elements is not due to indels in this region of MoTeR1. MoTeR2 products showed a similar pattern, with all the PCR products having a similar size demonstrating that size variation is not due to this region of MoTeR2. The *Bam*HI site was retained in all MoTeR1 PCR amplicons tested as two bands were seen after the PCR amplicons were digested with *Bam*HI (Figure 3-9-A, lanes 10-17). *Bam*HI cleaved most of the PCR products of MoTeR2. The exception to this is found in the isolate KS320 where three bands were observed following *Bam*HI digestion (Figure 3-9-A, lane 14). Two of these bands corresponded to the expected sizes of the digested PCR product, while the third band corresponded to the size of the undigested PCR amplicon.

3.3 Discussion

3.3.1 MoTeRs are members of a new class of non-LTR retrotransposons

One of the overall objectives of the experiments conducted was to characterize MoTeR1 and MoTeR2 structure. The general features of MoTeRs are outlined in Figure 3-1. MoTeR1 included a variable 5' end, an open reading frame for a reverse transcriptase (with a zinc finger domain, an RT domain, and an REL-endo domain), and a 3' end that is not adenylated post-transcriptionally. MoTeR2 had 5' and 3' ends identical to MoTeR1, but did not encode the reverse transcriptase. Since MoTeR2 lacks the reverse transcriptase, it may rely on MoTeR1 for its transcription. MoTeR2 may be similar to SINE-like elements. In the case of Mg-SINE the 3' end of the element is similar to the 3'end of the MGL retrotransposon (THON *et al.* 2004), and may use the MGL machinery for its retrotransposition (KACHROO *et al.* 1995). The MoTeR2 element, however, does not have a RNA polymerase III promoter typical of SINEs (GILBERT and LABUDA 1999). Because MoTeR2s are not likely to be transcribed by polymerase III they could represent a class of Short Internally Deleted Elements (SIDEs) similar to those found in the R2 retrotransposon system (EICKBUSH and EICKBUSH 2012).

Alternatively, the MoTeR2 could be involved in MoTeR1 transposition similar to the Het-A and TART non-LTRS of Drosophila. In this system transposition may only be possible if two separate non-LTRs cooperate. One of the non-LTRs (TART) encodes a RT, and the other element (Het-A) encodes a gag protein that efficiently targets telomeres (CASACUBERTA and PARDUE 2005).

The MoTeR1 element has all the necessary components to be an active retrotransposon, and it would be expected that MoTeR1 would be expressed. In a preliminary experiment expression of both MoTeRs was detected (Appendix B), but subsequent experiments did not give reliable results. This could be due to very low levels of expression in these isolates. Further work is needed to determine the expression profiles of MoTeRs. These results were exciting as MoTeRs may be the first example of an active telomeric non-LTR retroelement in fungi.

The close relationship of the RT and REL-endo in MoTeR1 to telomeric fungal retrotransposons (NhTER1, Cnl1, and FoNLR9) suggests that MoTeR1 is a member of a telomere specific clade of non-LTR retrotransposon. The MoTeR1, NhTeR1, Cnl1, and FoNLR9 were closely related to members of the CRE clade. Members of the CRE clade have a single ORF encoding a reverse transcriptase with a C-terminal endonuclease domain. CRE-like retrotransposons often exhibit target specificities for tandemly repeated DNA sequences such as rRNA genes (BURKE *et al.* 1995), spliced leader exons (AKSOY *et al.* 1990; VILLANUEVA *et al.* 1991), and, in the case of the Genie retrotransposon, repeats that are located near telomeres (BURKE *et al.* 2002).

According to the phylogenetic analysis, telomeric transposons appear to have evolved several times within the non-LTR retrotransposon clade (Figure 3-2). The TART and TAHRE elements in Drosophila and the SART and TRAS elements in *Bombyx mori* utilize different mechanisms to accomplish their insertion at chromosome ends. In these systems the non-LTR retrotransposon has taken on an additional telomere maintenance function due to lack of telomerase expression in

the case of *Drosophila melanogaster* (BIESSMANN and MASON 1997) and a weak telomerase expression in the case of *Bombyx mori* (OSANAI *et al.* 2006). Whether MoTeRs play a role in telomere maintenance is not known. These isolates could also possibly have weak expression of telomerase, and thus the MoTeRs could be involved in compensating for this weakness. However, the length of the telomeres in the GLS isolates of *M. oryzae* does not support this conclusion. A further study into the expression levels of telomerase in different *M. oryzae* isolates could help in elucidating at least part of this question. Some preliminary evidence using telomerase knockouts in MoTeR containing isolates have shown amplification of MoTeR sequences in telomerase-minus strains (Farman, unpublished data).

3.3.2 Variability of the 5' ends of MoTeRs

Another objective of the experiments conducted was to characterize MoTeR1 and MoTeR2 variability. MoTeRs are highly variable at their 5' ends, while being highly similar in other regions. The FH isolate showed two strong hybridization bands in the telomere blot corresponding to MoTeR1 and MoTeR2, while there are at least three known sizes of MoTeR1 in the FH genome. This suggested that in GLS isolates there is at least one MoTeR "type" present at most chromosome ends, and that the variant forms are at sub-terminal locations. Junction PCR detected at least four MoTeR variants in LpKY97-1A (Figure 3-8) at sub-terminal locations; only two of these were represented in the telomere Southern hybridization (Figure 3-7). LpKY97-1a(1) was of similar size to the full-length element in FH, but no terminal MoTeRs are similar in size between FH and LpKY97-1A in telomere Southern hybridizations suggesting that LpKY97-1a(1) may be a non-terminal MoTeR1.

Variation between 5' ends was observed between different GLS isolates. For example, extension of STR(C) within the MoTeRs was observed in CHW showing an increase in length of the element. Extensions of repeats could increase the size of the MoTeRs in a given isolate. It is not known whether these internal duplications would cause lower efficiency of transposition by increasing the likelihood of inefficient reverse transcription through premature disassociation of the RT at the repeated sequence.

The terminal MoTeRs of most of the GLS isolates analyzed harbor telomere variant STRs at their 5' ends [STR(A) (CCCAAA) and STR(B) (CCCGAA)]. In the Southern analysis of MoTeR1(5') the MoTeR1(5') probe failed to hybridize to the WK isolates suggesting that they had lost the 5' end of MoTeR1 and MoTeR2. The loss of the 5' end of an element could limit the ability of the element to transpose efficiently into telomeres. For example, the R2 non-LTR retrotransposon inserts specifically into the 28S RNA genes in diverse groups of eukaryotes (EICKBUSH and EICKBUSH 2012; KOJIMA and FUJIWARA 2004). The R2 reverse transcriptase binds a segment of the R2 RNA near the 5' end of the transcript. This causes a conformational change in the R2 reverse transcriptase, which allows it to bind to downstream of the genomic target. The binding of downstream DNA allows for more efficient transposition (CHRISTENSEN *et al.* 2006). Transposition was more precise in R2 retrotransposons with intact 5' ends (BURKE *et al.* 1999). The loss of the 5' end in MoTeRs could lead to instability of telomeres by inefficient transposition at the chromosome ends.

Long telomere repeat tracts were present in a few of the junctions between MoTeRs in RGNJ $(CCCTAA)_{18}$ and PL1 $(CCCTAA)_{28}$. These interstitial telomeres could possibly cause stalling of the replication fork during replication. This would lead to a double strand break (DSB), which has been observed at interstitial telomere repeats in yeast (IVESSA *et al.* 2002). This could lead to telomere instability in these isolates, and be a source for generating variability within these GLS isolates.

3.3.3 Possible endonuclease site of MoTeRs

When the MoTeRs occurred in tandem, they were always found in head to tail orientation similar to the Het-A, TAHRE and TART retrotransposons that make up the chromosome ends in *D. melanogaster* (ABAD *et al.* 2004b; MASON and BIESSMANN 1995; PARDUE and DEBARYSHE 2003). However, in the GLS isolates of *M. oryzae*, telomerase appears to be active with long hexanucleotide repeats of CCCTAA found at the chromosome ends. Analysis of MoTeR to MoTeR junctions showed that telomere sequences (CCCTAA) were commonly found at the border between two MoTeRs (Table 3-4). In Southern analysis of *Pst*I genomic digests the MoTeRs almost always cohybridized with telomeres (Figure 3-4). These experiments, along with previous data showing MoTeRs are bordered by telomere repeats (FARMAN 2007), suggest that MoTeRs insert into tandemly repeated telomere sequences. Characterization of *N. haematococca* and *C. neoformans* genomes showed that copies of this class of retrotransposons were also bordered by telomere repeats (data not shown).

The endonuclease of the MoTeR1 is likely targeting telomeric repeats, and may nick the target site between the A and C of the CCCTAA strand and between G

and T on the opposite strand. Other distantly related telomere specific non-LTR elements with APE endonucleases show site specificity similar to MoTeRs. In particular, the SART1 endonuclease digests between the A and T on the CCTAA strand, and at T and G on the TTAGG strand. TRAS, another member in this group, has an endonuclease that nicks between C and T on the CCTAA strand, and between A and T on the TTAGG strand (FUJIWARA *et al.* 2005). At least 2 units of the telomere repeat are needed for endonucleolytic activity of the TRAS1 endonuclease (ANZAI *et al.* 2001). The endonuclease of MoTeRs may not have strict sequence requirements. There were several cases where the CCCTAA sequence was not present in the MoTeR-to-MoTeR junctions, but telomere like STR(B) (CCCAAA) was present, which could suggest that a thymine in the 4th position may not be crucial for recognition by the endonuclease or that the target site is AACCC. Another possible explanation of the lack of telomere repeats could be a deletion at the insertion site, which has been noted in other non-LTR retrotransposons (BURKE *et al.* 1987; JAKUBCZAK *et al.* 1990). Further, work is necessary to determine the sequence specificity of the MoTeR1 REL endonuclease.

It is curious that the group of telomeric retrotransposons, in which MoTeRs belong, all have a substitution for the proline that is common in the PD(X_{12-14})D domain of the non-LTRs with REL type endonucleases. In MoTeR1, NhTeR1, and FoNLR9 alanine is substituted for proline. CRE2, a member of the CRE clade, has a substitution of the proline with a lysine in this domain. Evolutionary analysis of the REL type endonucleases indicate that they share similarities with Type II restriction endonucleases (EICKBUSH 2002). There are other suspected Type II endonucleases

which do not have the proline in the PD(X_{12-14})D domain (MENON *et al.* 2010) . The effect of proline substitutions on the activity of the endonuclease is not known. A site-directed mutagenesis study on the restriction enzyme of *Eco*RV in which the proline is changed to an alanine showed reduced activity (SELENT *et al.* 1992). Most research on the activity of this domain has focused on the aspartic acid residues, and their importance in proper endonuclease function. Specifically in R2Bm, mutations to the first aspartic acid residue of this domain disrupted the ability to cleave DNA, but did not affect the binding of the RT protein or target primed reverse transcription (YANG *et al.* 1999).

3.3.4 MoTeRs distribution

The last major objective of the experiments in this chapter was to determine the distribution of MoTeRs within the GLS isolates. Based on cohybridization in Southern analyses MoTeRs were associated with a majority of the telomeres in the GLS isolates. Out of the 136 telomere restriction fragments from various GLS isolates that were tallied, 123 cohybridize with MoTeR probes (Table 3-2). In the telomere restriction fragments that cohybridize to MoTeRs, they could have only MoTeR1 (42/123), only MoTeR2 (25/123), or both MoTeRs (56/123). This correlates to the different MoTeR1 and MoTeR2 configurations in the telomeres in FH (FARMAN 2007). The close association with telomeres makes the MoTeRs likely candidates for involvement in telomere restriction fragment variation.

The copy number of MoTeRs was variable among the different GLS isolates tested, and the GLS isolates have more copies on average of MoTeR1 (~16) than MoTeR2 (~11). CHRF in particular has a high number of MoTeR1 copies (~27) and

MoTeR2 copies (~23), while PL3 has a low number of MoTER1 copies (~11) and did not appear to have MoTeR2. It might be expected that isolates with more MoTeRs might show more telomere variation as there would be a greater chance of recombination between elements, a greater likelihood of inefficient DNA replication due to the interstitial telomere repeats between MoTeRs in arrays, and possibly more MoTeRs actively transposing into telomere repeats. Furthermore, the individual telomeres that harbor longer arrays of the MoTeRs might be less stable than telomeres with fewer MoTeRs. Further research into this area could examine the stability of individual telomeres through several mitotic generations.

The total number of MoTeRs that a telomere can effectively harbor and the mechanism by which MoTeR copy number can be controlled is unknown. Several different mechanisms to limit the copy number of transposable elements have been discovered in other TE systems: including methylation (NAKAYASHIKI *et al.* 2001a), transcriptional repression (PRUD'HOMME *et al.* 1995), homology-dependent gene silencing (JENSEN *et al.* 1999), RNAi (SAVITSKY *et al.* 2006), and overproduction inhibition (LOHE and HARTL 1996).

3.3.5 Telomere instability in subcultures

In LpKY97-1A a new telomere band appeared in a subculture indicating instability of its telomeres. This telomere fragment cohybridized with the MoTeR1 probe, but not the MoTeR5' probe (Figure 3-7). There are two likely explanations for the appearance of this de novo MoTeR1. One explanation is that a MoTeR1 became truncated past the 5' end probe, and then telomere sequence was added to the end by telomerase. The other possible explanation is that a truncated form of MoTeR1

transposed into a telomere sequence. Telomere instability in culture has been previously observed in another GLS isolate, FaMS96-1 (FARMAN 2002). Additionally, new MoTeRs bands appeared in the Southern analyses of subcultures of both RGNJ and LpKY97-1A indicating that MoTeRs may play a role in the instability of telomeres in GLS isolates (Figure 3-7 C and D). Further testing on the relative instability of telomere restriction fragments and MoTeR restriction fragments of GLS isolates in culture is needed to determine if this is a common phenomenon among GLS cultures. A more in-depth analysis could reveal that MoTeRs are the main cause behind the "hypervariability" observed in the telomeres of field isolates.

Table 3-1 Short Tandem Repeats within *Magnaporthe oryzae* Telomeric Retrotransposons.

Repeat	Consensus Sequence	Length	Location	Copy #
MoTeRs				
A	CCCGAA	6	1-9	1.5
B	CCCAAA	6	10-57	8
C	GGGGCTATTTATGCGCTTTAATTTGTGC	28	474-550	2.8
D	TATTT(ATGCGCTTTAATTT)GAGCGGGCTATTA(ATGCGCTTTAATTT)ACAAATTT	55	507-633	2.3
E	ATTGCTATTATTATCGTTACTATTATTATT	30	665-721	1.9
F	TTTCGTAGGCTTTGC	15	791-849	3.9
MoTeR1 only				
G	TTTACCTGTTTTATTAGCGGTTTACCTGCTTTTATTACCTGGTTCCCC	48	906-1171	5.5
H	GTTTTTACTAGCAGTTAAATTTACCTTTTTAAGGTTATTTACCTGCTTTTATTCACAGGGCACCCCT	67	1200-1386	2.8
I	ATAACCCCAGGGTTA	15	1901-1950	3.3
J	TTTTTG	6	4687-4829	23.8
K	TTTTTTC	7	4913-4938	3.7
MoTeR2 only				
L	TTTTACCTGCTT	12	857-882	2.2
M	GACCCCGACCCTGACCCG	18	982-1019	2.1
N	TATTTATTCGATT	13	1245-1271	2.1
O	ATTGGATTAATCCGCTAAATAA	22	1516-1572	2.6
P	TTTTTTCTTTTTC	13	1612-1636	1.9

Sequences within parentheses are duplicated within a single repeat. Sequences that are underlined are present in two different repeats.

Table 3-2. MoTeR cohybridization with telomeric fragments in GLS isolates

Isolate	# Telomere Band fragments Counted	MoTER1	MoTER2	Both MoTeRs	Telomere Only
PL1-1	11	1	0	7	3
PL2-1	9	1	3	4	1
PL3-1	13	13	0	0	0
CHW	10	3	3	4	0
CHRF	8	2	2	3	1
GG-PW9	7	1	2	4	0
GG13	9	2	1	6	0
KS320	12	2	3	5	2
KS330	11	1	3	5	2
LPOH97-1	10	4	2	4	0
LPOH3	9	4	2	2	1
RGNJ	10	2	2	5	1
TFMS	12	5	1	4	2
TFRGA	5	1	1	3	0
Total	136	42	25	56	13
Average Per Isolate	10.5	3.2	1.9	4.3	1.0

Table 3-3. Copy number analyses of telomere and MoTeRs using densitometric scanning and visual estimation.

Isolate	**Densitometric Scan**			**Visual Estimation**		
	Telomere	MoTER1	MoTER2	Telomere	MoTER1	MoTER2
PL1	14.91	19.29	12.91	13	14	13
PL2	15.03	22.41	12.98	14	14	12
PL3	14.68	14.80	0	15	15	0
CHW	9.61	10.52	9.12	10	10	9
CHRF	13.91	27.37	22.86	14	14	15
GG-9	14.79	9.97	13.36	12	7	11
GG-13	12.93	15.62	10.05	13	11	11
KS320	14.21	14.51	12.21	14	10	10
KS330	13.85	12.27	19.03	14	12	13
LPOH97-1	15.51	15.93	11.20	15	16	12
LPOH3	13.82	15.54	6.01	14	16	5
RGNJ	17.29	15.64	14.73	15	16	13
TFMS	14.26	12.44	4.21	14	11	5
TFRGA	15.32	11.04	9.58	9	8	10
TOTAL	200.12	217.35	158.25	186	174	139
Average	14.29	15.53	11.30	13.29	12.43	9.93

Table 3-4 5' end variation in MoTeR-to-MoTeR junctions

Junction Type	...AATAAAGCGCGAATTAAAA----------to 5' MoTeR	Isolate
MoTeR2(3') to MoTeR(5')	...(CCCTAA)$_4$[CCCAAA](CCCTAA)[CCCAAA]$_{10}$CCCGAA...	LpKY97-1A
	...(CCCTAA)$_3$[CCCAAA](CCCTAA)[CCCAAA]$_{10}$CCCGAA..	LpKY97-1A
	...(CCCTAA)$_2$(CCCGAA)$_2$[CCCAAA]$_8$CCCGAA	FH, RGNJ
	...(CCCTAA)[CCCAAA]$_8$CCCGAA...	FH
	...(CCCTAA)$_3$[CCCAAA]$_7$CCCGAA...	FH, LpKY97-1A
	...(CCCTAA)(CCCGAA)[CCCAAA]$_6$CCCGAA...	FH, RGNJ
	...(CCCGAA)[CCCAAA]$_6$CCCGAA...	RGNJ
	...(CCCTAA)[CCCAAA]$_6$CCCGAA...	PL1, RGNJ
	...(CCCTAA)$_3$[CCCAAA]$_4$CCCGAA...	FH, RGNJ
	...[CCCAAA]$_3$CCCGAA...	PL1, KS320
	...(CCCTAA)$_{28}$[CCCAAA]$_2$CCCGAA...	PL1
	...(CCCTAA)$_2$[CCCAAA]$_2$CCCGAA...	PL1
MoTeR1(3') to MoTeR(5')	...(CCCTAA)$_3$[CCCAAA]$_{11}$CCCGAA...	LpKY97-1A
	...(CCCTAA)$_3$[CCCAAA](CCTAA)$_2$[CCCAAA]$_{10}$CCCGAA...	LpKY97-1A
	...(CCCTAA)$_2$[CCCAAA](CCCTAA)[CCCAAA]$_{10}$CCCGAA...	LpKY97-1A
	...(CCCTAA)(CCCGAA)$_2$[CCCAAA]$_8$CCCGAA...	FH
	...(CCCTAA)$_2$-**MoTeR-3'(11 bp)-(CCCGAA)$_2$**[CCCGAA]$_8$CCCGAA....	FH
	...(CCCTAA)(CCCGAA)[CCCAAA]$_6$CCCGAA...	FH
	...(CCCTAA)CCC-MoTeR-3'(40 bp)-(CCCTAA)$_5$(CCCGAA)[CCCAAA]$_5$CCCGAA...	RGNJ
	...(CCCTAA)CCCTAG(CCCTAA)$_{27}$(CCCGAA)(CCCTAA)(CCCGAA)[CCCAAA]$_5$CCCGAA...	RGNJ
	...(CCCTAA)$_{28}$(CCCGAA)(CCCTAA)(CCCGAA)[CCCAAA]$_5$CCCGAA...	RGNJ
	*..(CCCTAA)(GCCAAA)[CCCAAA]$_5$CCCGAA...	RGNJ
	...(CCCTAA)$_{18}$[CCCAAA]$_4$CCCGAA...	RGNJ
MoTeR1-MoTeR1	...(CCCTAA)$_5$(CCCGAA)[CCCAAA]$_4$CCCGAA...	FH
	...(CCCTAA)[CCCAAA]$_3$CCCGAA...	FH
MoTeR2-MoTeR1	...(CCCTAA)$_2$CCC-MoTeR1-5'(448 bp)...	FH
	...CCC-MoTeR1-5'(448 bp)...	FH
	...(CCCTAA)G-MoTeR1-5'(785)	RGNJ, FH, LpKY97-1A
	...(CCCTAA)$_3$-MoTER1-5'(364)	LpKY97-1A
	...(CCCTAA)$_3$[CCCTAA]$_7$CCCGAA...	LpKY97-1A
	...(CCCTAA)[CCCAAA](CCCTAA)[CCCAAA]$_{10}$CCCGAA...	LpKY97-1A

* indicates only one clone contained the junction. Bold text indicates a junction without the telomere repeat CCCTAA. Red letters indicate an uncommon mismatch in the telomere or telomere-like repeats. Number of base pairs in parenthesis indicates the truncation point of an element when it occurs after 5', and parenthesis indicates the number of base pairs left of the element when it occurs before MoTeR3'

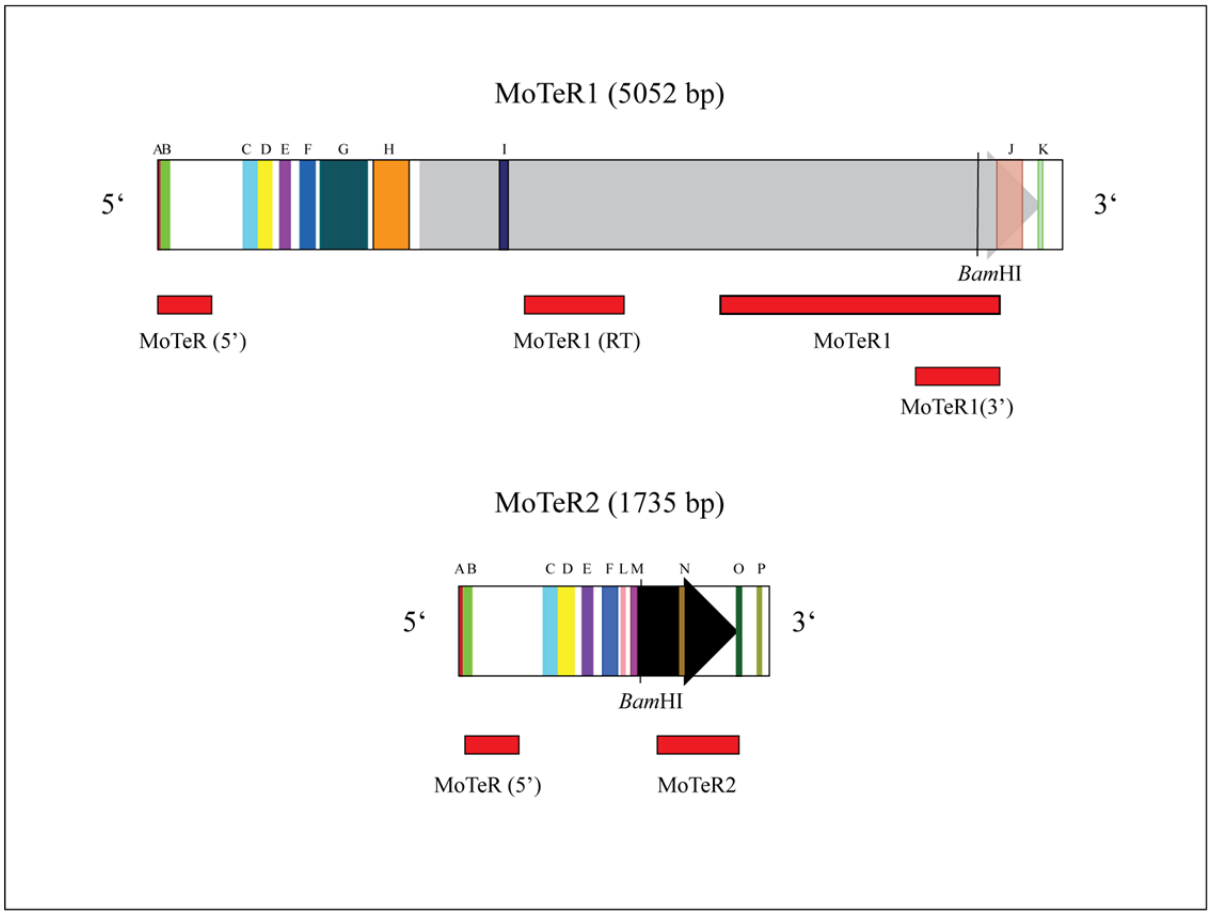

Figure 3-1. Structure of *Magnaporthe oryzae* Telomeric Retrotransposons (MoTeRs). The figure depicts the graphical representation of MoTeRs. The letters above the elements are different short tandem repeats (STR). The sequences of the STRs are available in Table 3-1. The total size of the STR is depicted in different shaded or colored boxes. The arrow represents the largest open reading frame within the MoTeR element. Red boxes indicate position in the MoTeR where the respective probe would bind. MoTeR1 is 5 kb in length and harbors eleven different repeats. MoTeR2 is 1.7 kb in length and harbors eleven different repeats.

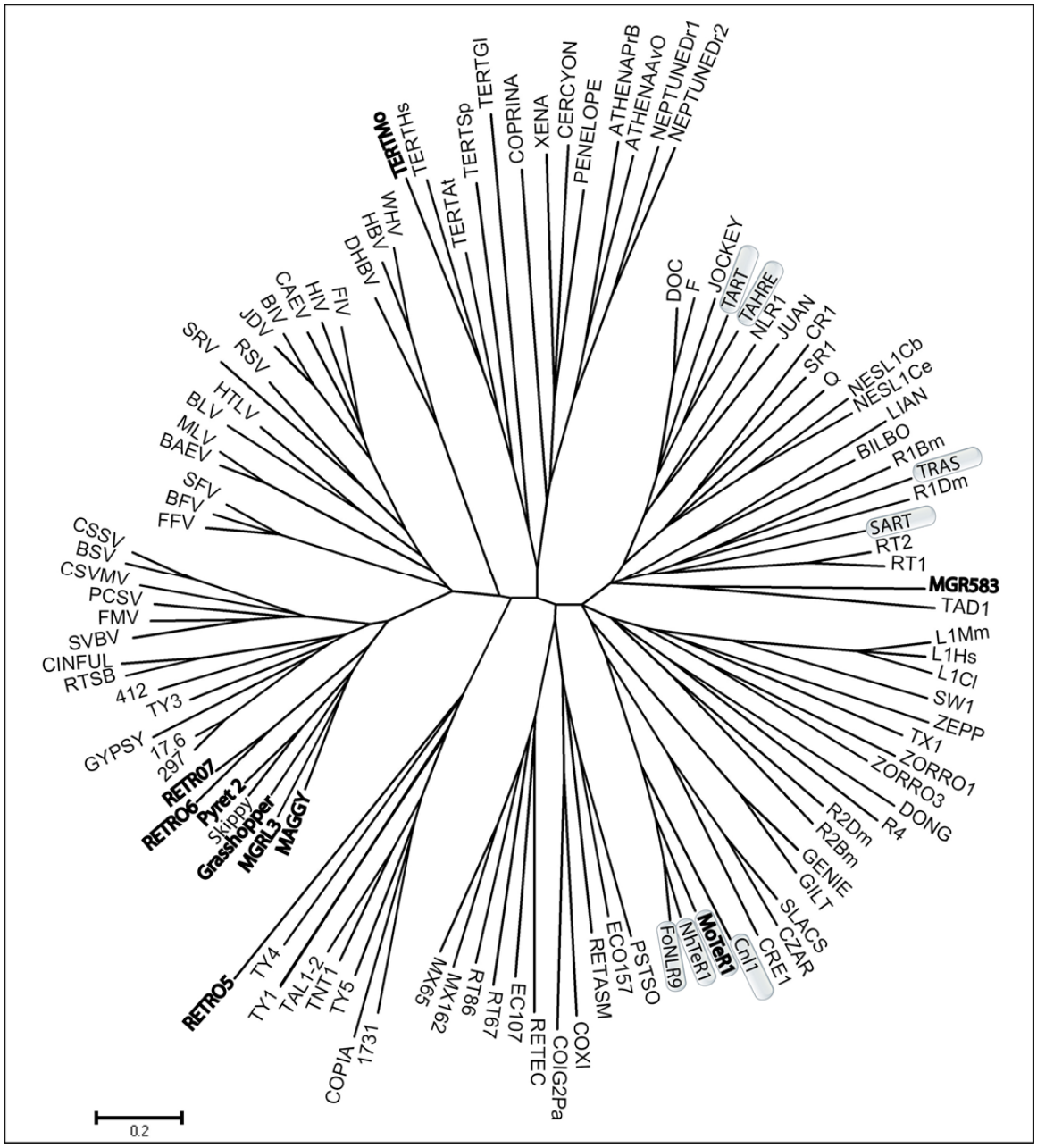

Figure 3-2. Phylogenetic relationship of MoTeR1 to other reverse transcriptases. Bold text indicates elements present in *Magnaporthe oryzae*. Gray ovals indicate the telomeric non-LTR retrotransposons. There were a total of 485 positions in the final data set, consisting of conserved RT domains. Positions with gaps or missing data were only eliminated in pairwise comparison. The phylogenetic tree was constructed using MEGA4 (TAMURA *et al.* 2007), by the Neighbor-joining method (SAITOU and NEI 1987). Poisson correction was used to calculate the evolutionary distance. The bootstrap consensus tree is inferred from 1000 replicates (FELSENSTEIN 1985), and branches with less than 50% bootstrap replicates are collapsed.

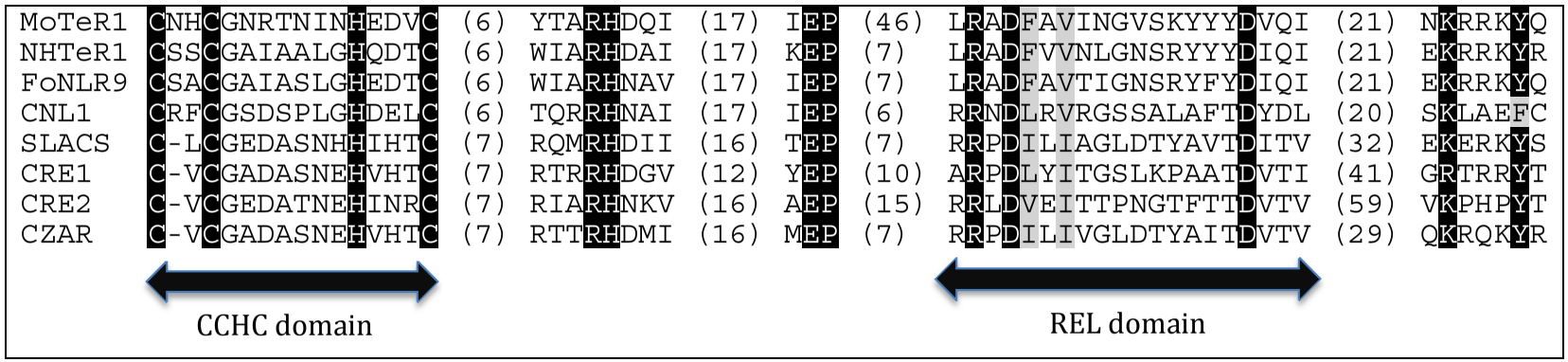

Figure 3-3. Comparison of the restriction enzyme-like (REL) endonuclease domains of MoTeR1 with closely related non-LTR retrotransposons. Highly conserved residues are highlighted in black, and semi-conservative substitutions are highlighted in gray. The number of amino acids between the conserved motifs is given in parenthesis. Modified from (VOLFF *et al.* 2001).

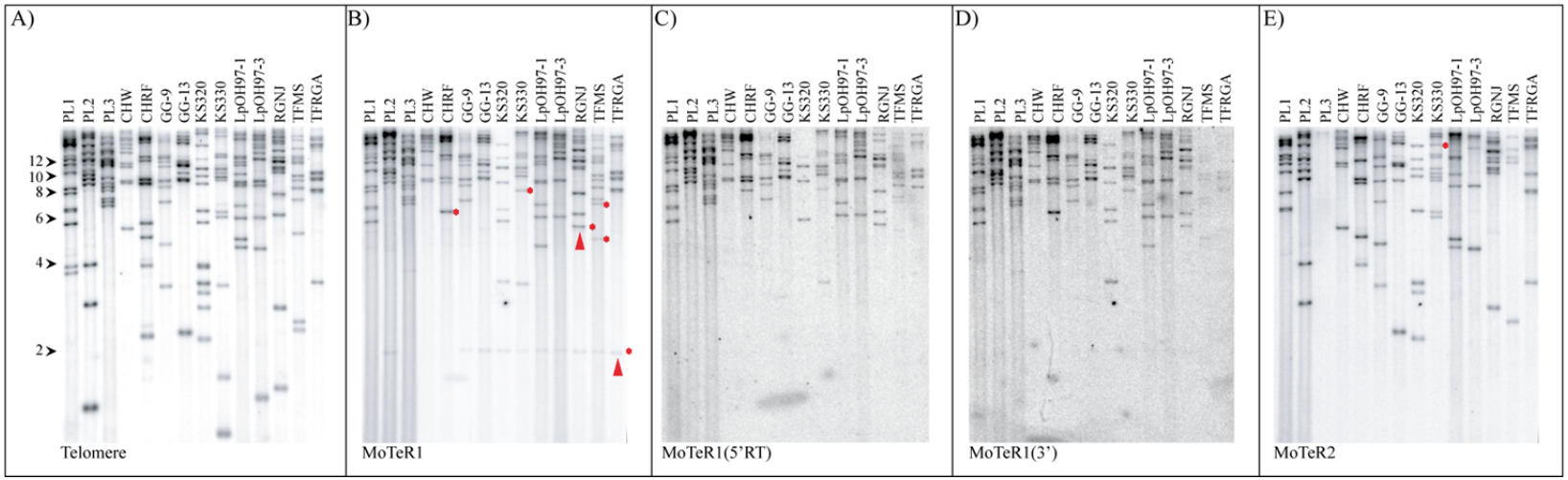

Figure 3-4. Southern blot analyses of genomic DNA from GLS isolates of *M. oryzae*. Genomic DNA was digested with *Pst*I, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized sequentially with A) Telomere, B) MoTeR1, E) MoTeR2, D) MoTeR1(5'RT), and E) MoTeR1(3'). Between hybridizations the blot was stripped. Black arrows indicate the size markers in kilobases. The asterisk (*) in 4-B and 4E indicates MoTeRs that did not cohybridize with telomere fragments. Data from this experiment had priority in band counting included in the tables. Red arrows in 4-B indicate MoTeR1 bands that did not cohybridize with telomere bands that were cloned.

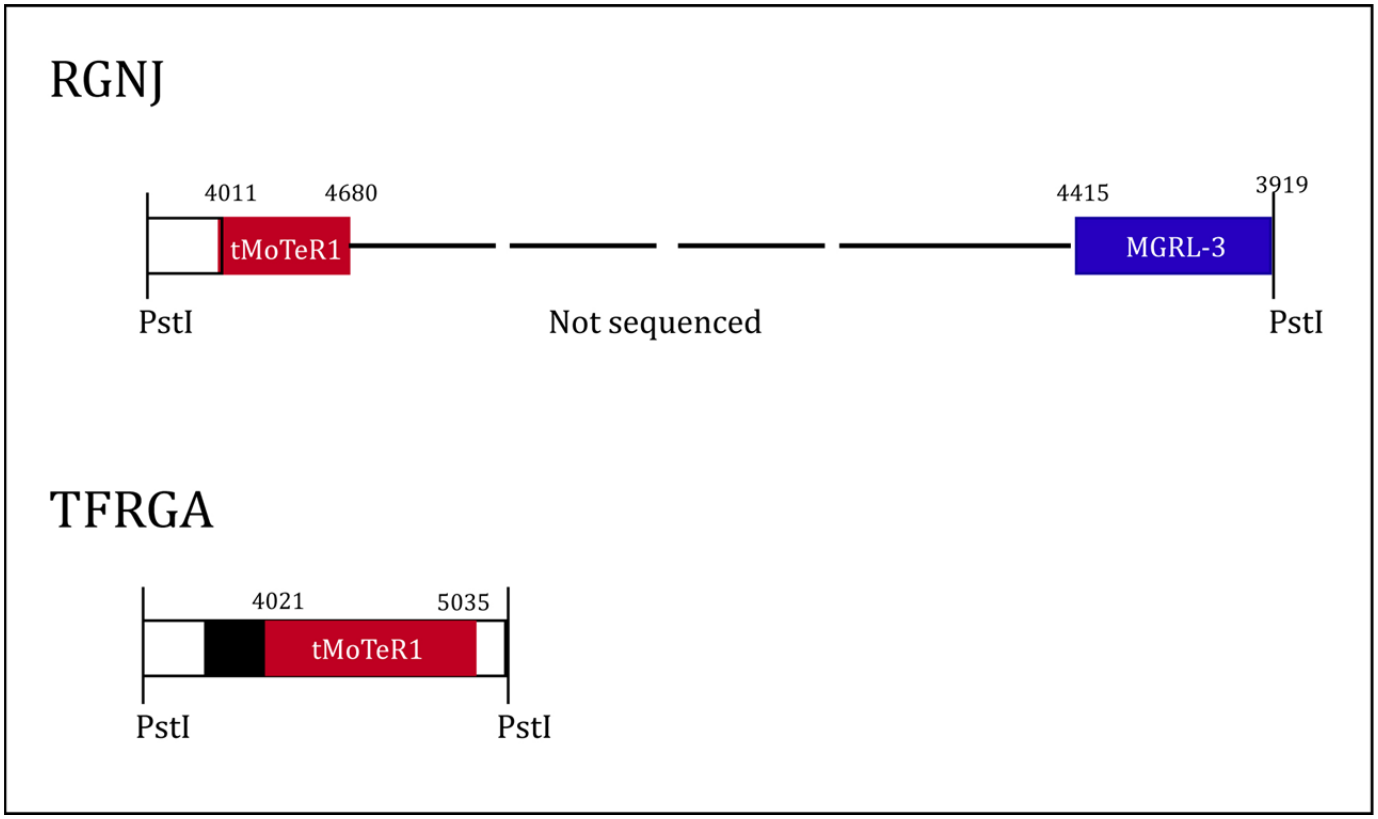

Figure 3-5. Internal MoTeR1 fragments cloned from *Pst*I digests of genomic DNA from RGNJ and TFRGA. PstI sites bordering the fragments are marked. Numbers indicate the positions in the elements where significant similarity was found in the local Blast analysis with the sequences of the fragments. A dashed line indicates portions of the fragment that were not sequenced. Red boxes labeled tMoTeR represent MoTeR1s truncated at different locations. The LTR retrotransposon MGRL-3 is labeled with a blue box around the portion that was sequenced in the experiment. The black box in TFRGA has some sequence similarity based on Blastn analysis to the telomeric clone 72H05 (position 8024-8339) located on telomere 12 in 70-15. White boxes with black borders are sequences with no significant similarity to the genome sequence of *M. oryzae*.

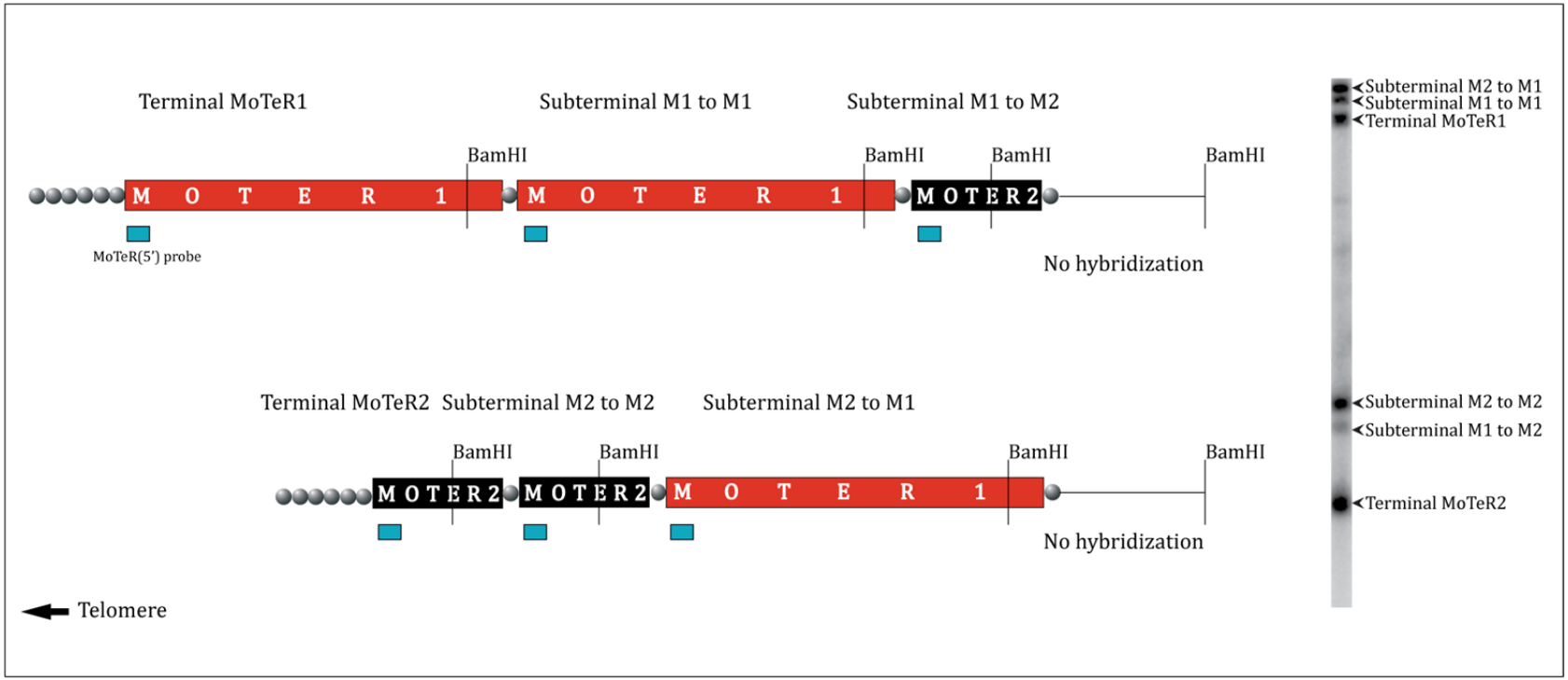

Figure 3-6. Expected composition of hybridization signals in the Southern hybridizations of *Bam*HI digests probed with MoTeR(5') in the gray leaf spot isolate FH. Representation of the possible combinations of MoTeR1 and MoTeR2 at the chromosome ends. Genomic DNA from isolate FH was digested with BamHI, electrophoresed on a 0.7% agarose gel, electroblotted to a nylon membrane, and hybridized with a MoTeR(5') probe. The membrane was then exposed to a phosphorimaging screen and the resulting phosphorimage is shown. Black boxes represent MoTeR2. Red boxes represent MoTeR1. Gray circles represent telomere repeats (TTAGGG). Blue boxes indicate the MoTeR(5') probe.

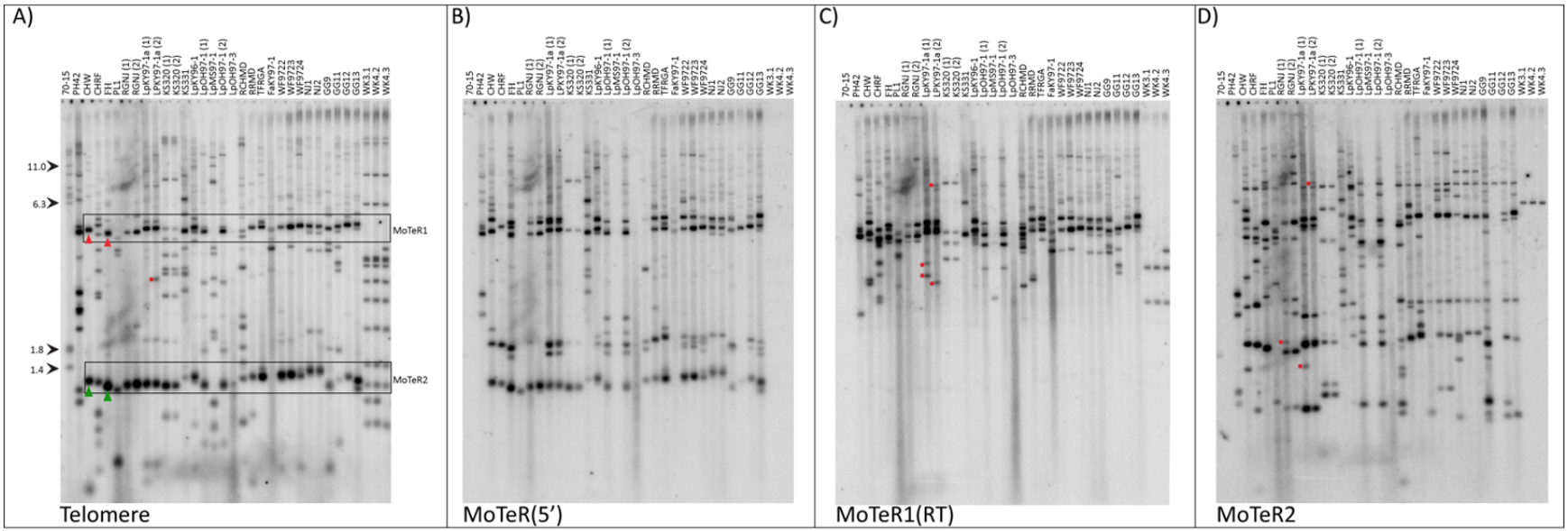

Figure 3-7. Southern analyses of genomic DNA from various GLS isolates. Genomic DNA was digested with *Bam*HI, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized sequentially with A) Telomere, B) MoTeR(5'), C) MoTeR1(RT), and D) MoTeR2. The membrane was then exposed to a phosphorimaging screen and the resulting phosphorimages are shown. Between hybridizations blots were stripped. Black arrows indicate the size markers in kilobases. The isolates are listed above the lanes, and numbers in parenthesis indicate an additional lab culture analyzed. In lane 1 is the laboratory rice infecting strain 70-15. In lane 2, PH42 is an *Eleusine indicia* host specialized isolate. In lanes 3-34, different GLS isolates are shown. Red asterisks show alterations in internal MoTeR fragments between different cultures of the same isolates. Red arrows indicate that the terminal MoTeR1 in CHW and FH have a different size. Green arrows indicate that the terminal MoTeR2 fragments in CHW and FH have a different size.

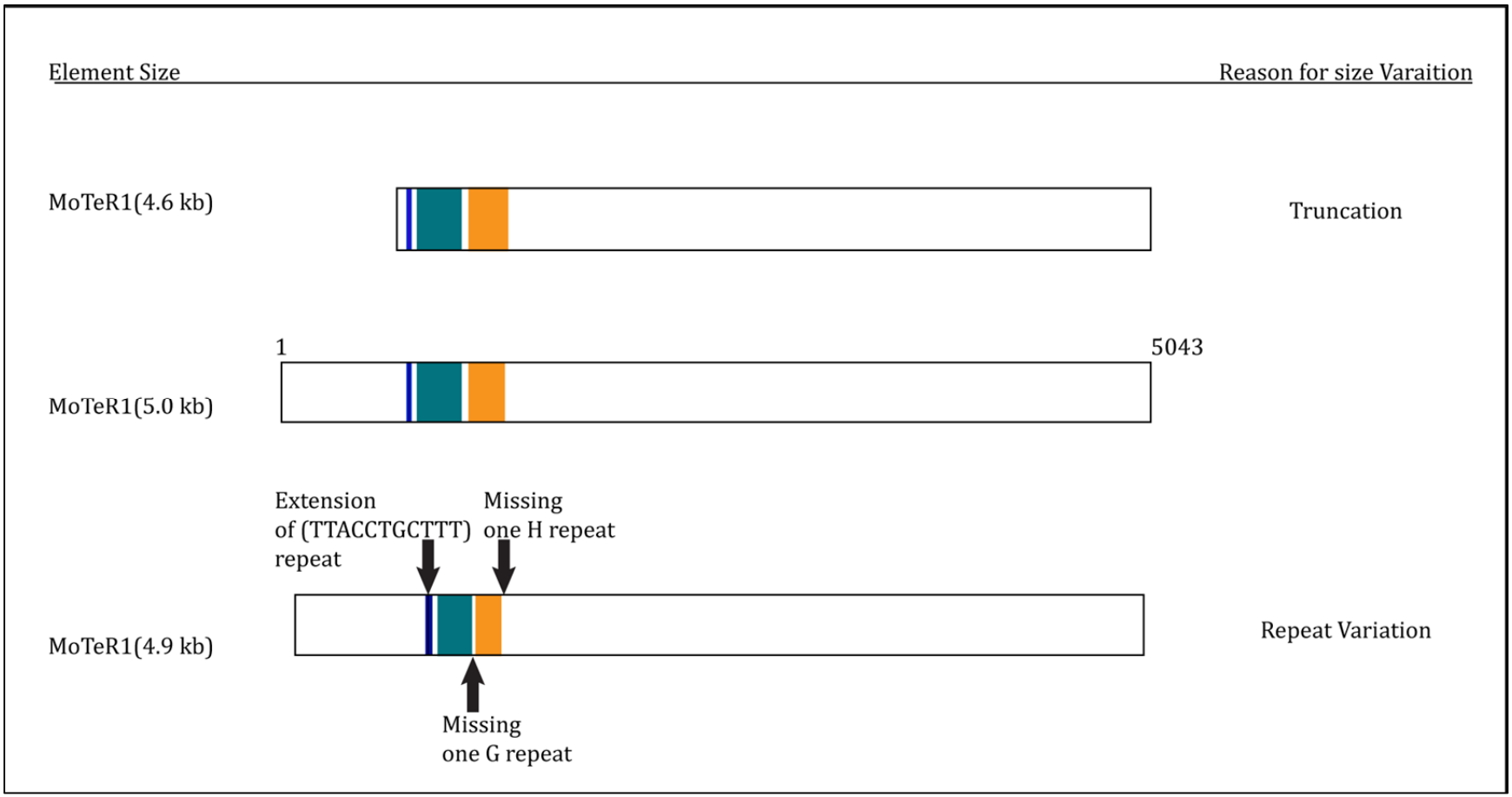

Figure 3-8. Graphical representation of the three different MoTeR1 elements in the FH isolate. The different filled boxes represent repeat units. Boxes filled in the same color represent the same repeat. The teal boxes represent the G repeat. The orange filled boxes represent the H repeat. The blue box represents a new repeated sequence (TTACCTGCTTT) that was seen in the MoTeR1 (4.9). The remaining white box outlining the MoTeR1 elements represents alignments that were nearly identical.

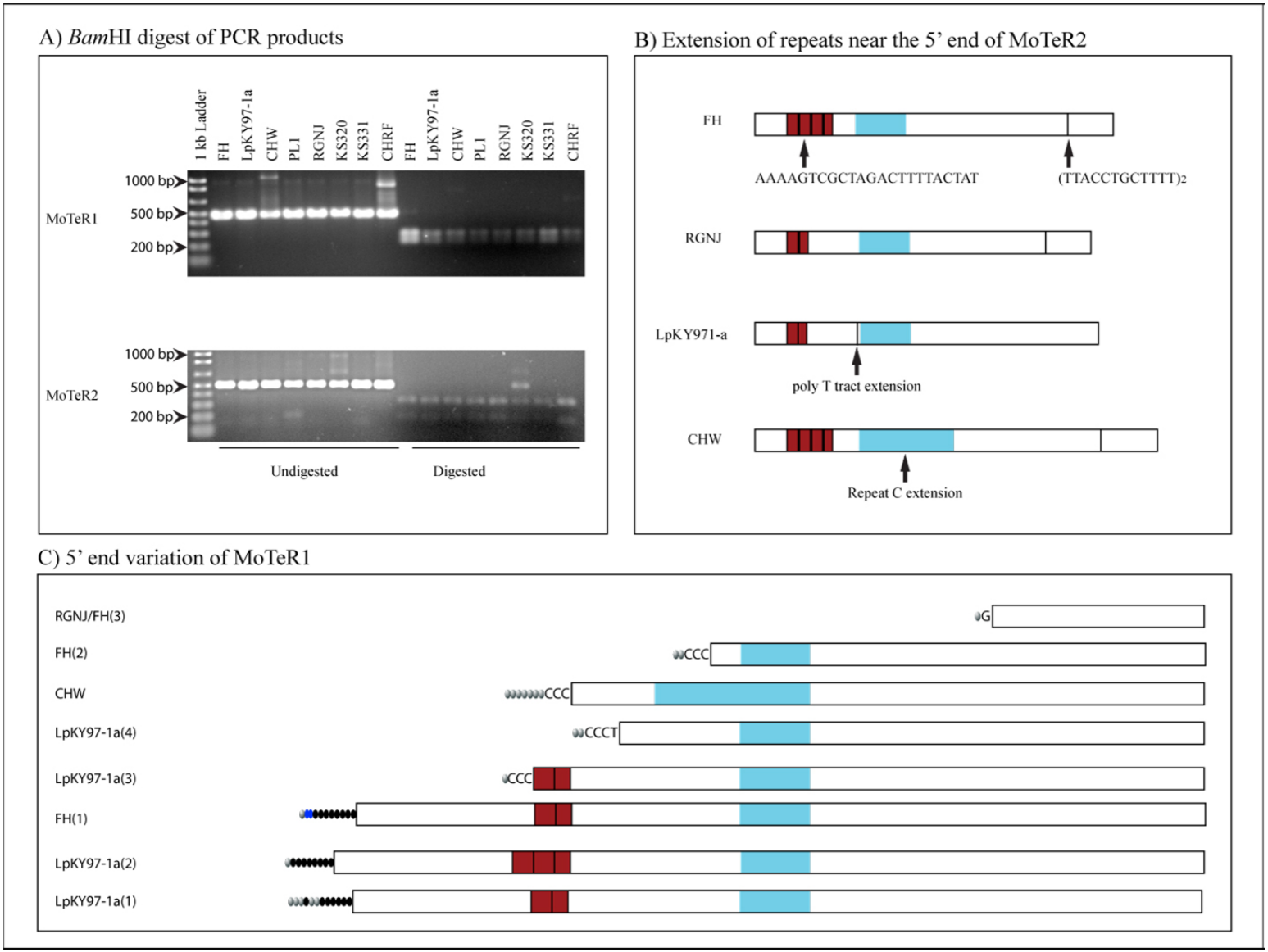
A) BamHI digest of PCR products
1 kb Ladder
FH
LpKY97-1a
CHW
PL1
RGNJ
KS320
KS331
CHRF
FH
LpKY97-1a
CHW
PL1
RGNJ
KS320
KS331
CHRF
1000 bp
500 bp
200 bp
MoTeR1
1000 bp
500 bp
200 bp
MoTeR2
Undigested
Digested
B) Extension of repeats near the 5' end of MoTeR2
FH
AAAAGTCGCTAGACTTTTACTAT
(TTACCTGCTTTT)2
RGNJ
LpKY971-a
poly T tract extension
CHW
Repeat C extension
C) 5' end variation of MoTeR1
RGNJ/FH(3)
G
FH(2)
CCC
CHW
CCC
LpKY97-1a(4)
CCCT
LpKY97-1a(3)
CCC
FH(1)
LpKY97-1a(2)
LpKY97-1a(1)

Figure 3-9. Characterization of size variation of MoTeRs in GLS isolates.
The figure shows the result for different possible reasons for band variation in Southern hybridizations. A) Shows the PCR amplification product in both MoTeR1 (top panel) and MoTeR2 (bottom panel) that spans the BamHI site in different GLS isolates (lanes 2-9). The PCR product was then digested with *Bam*HI to confirm presence of the BamHI site in the PCR product (lanes 10-17). B) Shows the graphical representation of alignments of sequence amplified by PCR from the 5' region of MoTeR2. Boxes with different patterns indicate different repeated elements. The light blue color represents repeat C, while the red colored boxes represents the sequence AAAAGTCGCTAGACTTTTACTAT. The remaining white box outlining the MoTeR2 5' sequence represents alignments that were nearly identical. C) Shows the graphical representation of the alignments of 5' end sequence. Numbers in parenthesis indicate the different 5' ends obtained in a single isolate. Gray circles indicate the telomere repeat TTAGGG. Blue circles represent a variant telomere repeat of TTCGGG, and black circles represent the variant telomere repeat TTTGGG. Telomere-like repeats not drawn to scale. Boxes containing colors represent different repeated elements. The light blue color represents repeat C, while the red colored boxes represents the sequence AAAAGTCGCTAGACTTTTACTAT. The remaining white box outlining the MoTeR1 5' end represents alignments that were nearly identical.

CHAPTER FOUR

Generation of New Telomeric Profiles in Mitotic Progeny

4.1 Introduction

DNA fingerprinting analysis has been used extensively to understand the population structure of (FARMAN 2002; LEVY *et al.* 1991; TANAKA and NAKAYASHIKI 2009; TOSA *et al.* 2007; VIJI *et al.* 2001) and the phylogenetic relationships among (KATO *et al.* 2000) different host specific isolates of *M. oryzae*. However, the use of different probes sometimes gives incongruent results. For example, gray leaf spot (GLS) field isolates collected from perennial ryegrass exhibit extreme telomere variability even when little variation exists at internal DNA loci (FARMAN and KIM 2005).

Several factors may lead to restriction fragment length polymorphisms in homologous fragments between different individuals. Point mutations could cause a loss or gain of a restriction site leading to new fingerprint patterns. The insertion or deletion of sequences between two restriction sites may lead to larger and smaller sized fragments, respectively. The insertion or deletion of an element with a restriction site will lead to restriction fragment length polymorphisms. There may also be more extensive rearrangements of the DNA (GRANT and SHOEMAKER 1997). Transposable elements play a role in generating restriction fragment length polymorphisms (LLOYD *et al.* 1987). They can insert into a new location adding new restriction sites and cause the loss of restriction sites if they transpose out of a location within the genome (BENDER *et al.* 1983; COLLINS and RUBIN 1982; DORING and STARLINGER 1984; KUFF *et al.* 1983). Major changes in restriction length fragments

can occur by chromosomal deletion resulting from ectopic recombination among transposable elements (KUPIEC and PETES 1988).

The *M. oryzae* genome has Class I and Class II transposable elements (DEAN *et al.* 2005). The Class I elements transpose through the use of an RNA intermediate, and can be divided into three types (LTRs, LINES, and SINES) (KEMPKEN and KUCK 1998). LTRs have long-terminal repeats bordering the ends of the elements and encode a reverse transcriptase. The LINES do not have the long-terminal repeats, but do encode a reverse transcriptase. The SINEs do not encode a reverse transcriptase. The known LTRs retrotransposons in *M. oryzae* include MAGGY (FARMAN *et al.* 1996b), Grasshopper (DOBINSON *et al.* 1993), MGLR-3 (KANG 2001), Inago1 (SANCHEZ *et al.* 2011), Pyret (NAKAYASHIKI *et al.* 2001b), RETRO6 (DEAN *et al.* 2005), RETRO7(DEAN *et al.* 2005), and RETRO5 (FARMAN *et al.* 2002). There are two known LINE-like elements in *M. oryzae*, MGL (MGR583) (HAMER *et al.* 1989a; KACHROO *et al.* 1997), and MoTeRs (FARMAN 2007). The SINE-like element in the genome of *M. oryzae* is Mg-SINE (KACHROO *et al.* 1995). MINE is a chimeric element between a MGL and WEIRD (FUDAL *et al.* 2005).

Class II elements are excised from the genome and transpose from one place in the genome to another (FINNEGAN 1989). There are four Class II transposable elements in *M. oryzae* including Pot2 (KACHROO *et al.* 1994), Pot3(MGR586) (FARMAN *et al.* 1996a), Pot4 (DEAN *et al.* 2005), and Occan (KITO *et al.* 2003).

In Southern analysis from the previous chapter, changes of telomere restriction profiles in separate cultures of the same GLS isolate were observed (Figure 3-7-A). This suggested that mutations during mitotic growth could lead to

the telomere variability. By comparison, rice blast (RB) isolates do not have extreme telomere variability. When internal DNA profiles among RB isolates are similar, the telomere profiles will also show high similarity (FARMAN 2007). DNA fingerprints using a variety of probes (MGR586, MAGGY, Pot2, MGL, MG-SINE) in the Rice Blast isolate 70-15 are highly stable even after ten mitotic generations, though a subtelomeric probe (TLH1) showed some instability (PARK *et al.* 2010). However, high mutation rates have been observed at particular loci in RB isolates (CHUMLEY and VALENT 1990; HAMER *et al.* 1989b), and genes involved in host-specificity have shown instability when located near chromosome ends (ZHOU *et al.* 2007).

Preliminary experiments indicated that there was a difference in relative stability of DNA fingerprints in mitotic progeny between a RB isolate (70-15) and a GLS isolate (LpKY97-1A) *in planta* and in culture. The objective of the experiments in this chapter was to compare the relative stability of DNA fingerprints in mitotic progeny of a RB isolate and a GLS isolate. The stability of telomeres was tested using a telomere-specific probe, and the internal chromosomal stability was tested using a probe from the Pot2 transposable element, which is distributed throughout the genome. If more chromosomal instability exists in telomeric or internal loci in the GLS isolate, one would expect its mitotic progeny to show more variation than the RB isolate's mitotic progeny.

4.2 Results

4.2.1 Rates of telomere change in a GLS isolate versus a RB isolate

Preliminary experiments suggested that there was a difference in telomere restriction fragment length stability in mitotic progeny between a rice blast isolate

and a GLS isolate (data not shown). To test whether the telomeres of a GLS isolate were less stable than those of a RB isolate, I monitored changes in telomere restriction profiles after single spore isolation of the initial culture in a RB isolate (70-15) and a GLS isolate (LpKY97-1A). Using single spores allows for removal of genetic variation that could have arisen during prior culturing. The purpose of the experimental design was to generate a total of three generations of single spore isolates [one generation of spores that made up the initial inoculum for plant infection (G1), and two generations of single spores recovered from lesions (G2 and G3)] from two different host-specific types [LpKY97-1A (GLS isolate) and 70-15 (a rice infecting isolate)].

After generating the initial single spore cultures in 70-15 and LpKY97-1A (known as the G0 cultures), conidia were collected by methods described in Chapter 2. A total of 40 monoconidial isolates were collected from the two fungal strains (20 from 70-15 and 20 from LpKY97-1A) and labeled as first generation (G1) progeny. The remaining conidial suspensions were used for plant infection as outlined below. The rice (*Oryza sativa*) cultivar 51583 was inoculated with the 70-15 conidial suspension and the perennial ryegrass (*Lolium perenne*) cultivar Linn was inoculated with the LpKY97-1A. Leaves showing lesions were clipped and placed in a moist chamber to allow for sporulation of the fungus. After 3-5 d in the moist chamber, spores were collected. Forty total single spores were collected (20 from LpKY97-1A and 20 from 70-15) and cultures were designated as the second generation (G2) progeny. The remaining suspensions were used to inoculate a second set of plants. Inoculation, plant growth, and spore collection was repeated as

described above. Then 40 more single spores were collected (20 from LpKY97-1A and 20 from 70-15) and designated as the third generation (G3) progeny.

A total of 120 single spore progeny were collected over three generations in two different *M. oryzae* isolates (60 from LpKY97-1A and 60 from 70-15). Of the 40 single spores collected at each spore generation, DNA from 38 of the mitotic progeny was extracted, digested with *Pst*I, electrophoresed, and examined by Southern hybridization with a telomere probe. *Pst*I was chosen due to the lack of a restriction site within the MoTeRs. The resulting telomere fragments from genomic DNA digested with *Pst*I would include the telomere through the MoTeRs if present, and to the first *Pst*I site at the chromosome end. Thus a potential change in the length of a MoTeR array at a telomere could be detected in the Southern analysis. A contraction of a MoTeR array would lead to shorter fragments, while an expansion in a MoTeR array would lead to longer fragments. Each hybridizing fragment in the Southern blot probed with telomere was counted as a single telomere regardless of intensity. The appearance of a new band in the DNA fingerprint of mitotic progeny that was absent in the parental isolate was counted as a gain, while the absence of a band in the DNA fingerprint of mitotic progeny that was present in the parental isolate was counted as a loss.

The Southern analyses with the telomere probe of *Pst*I digested DNA from mitotic progeny of 70-15 revealed that the telomeres were stable (Figure 4-1). No changes in the telomere restriction profiles were observed in mitotic progeny produced during culturing on oatmeal agar plates (Figure 4-1-A). Only one difference in the telomere restriction profile was observed in a mitotic progeny of

70-15 following two disease cycles in the plant (marked with an asterisk Figure 4-1-B, lane 6). This band had a faint hybridization signal, which implies that it had not become fixed in the culture.

In Southern analyses of *Pst*I digested DNA from LpKY97-1A, a higher telomere instability was observed among mitotic progeny (Figure 4-1-C and D) than was previously observed in the mitotic progeny of 70-15 (4-1 A and B). By the third generation none of the mitotic progeny of LpKY97-1A shared the same telomere fingerprint as the original starting culture. There were six band gains (.32 per progeny) in G1, 35 band gains (1.89 per progeny) in G2, and 50 band gains (2.6 per progeny) in G3 (Table 4-1). Based on post hoc t test of pairwise One-way ANOVA, there were significantly more gains and losses in LpKY97-1A mitotic progeny collected from plants than in mitotic progeny collected from cultures (*p-value* <0.001).

4.2.2 Telomere fragment variation in mitotic progeny of other GLS isolates

There was a difference in telomere stability between 70-15 and LpKY97-1A, but it was unknown whether this was a general phenomenon within GLS isolates. To determine if the instability of telomeres was common among Lolium pathotypes, two additional isolates (KS320 and RGNJ) were tested by Southern analysis with a telomere probe. The general outline of the methods used in this experiment is described below. KS320 and RGNJ were genetically purified using the single-sporing method. The mycelium was allowed to grow across an oatmeal agar plate, and the spores were harvested using the same methods described in Chapter 2. Twenty germinated mitotic progeny were collected, cultured, and their DNA was

extracted. *Pst*I-digested DNA from 19 single spores were analyzed by Southern hybridization with a telomere probe (Figure 4-2). Band gains and losses in mitotic progeny were tabulated (Table 4-2). Two-sample t-tests were used to compare differences in band gains and losses in the RB isolate (70-15) and the GLS isolates (LpKY97-1A, RGNJ, and KS320).

In KS320, three of the 19 mitotic progeny (16%) analyzed showed a different telomere restriction profile relative to the initial isolate (Figure 4-2). In RGNJ, mitotic progeny showed a different telomere restriction profile in five out 19 progeny analyzed (26%). This was a higher number than was seen in KS320 but lower than in LpKY97-1A (31%). LpKY97-1A and RGNJ had significantly more band gains (p-*value* < 0.01 and p-*value* < 0.05 respectively) than 70-15 (Table 4-2). However, KS320 was relatively stable by comparison, and did not have a significantly higher number of band changes than 70-15 (p-*value* < 0.15). No significant difference between the isolates was observed in telomere band losses.

4.2.3 Comparison of telomere and Pot2 profile variation

A previous study of GLS isolates had indicated that telomere profiles were more variable than internal chromosomal locations (FARMAN and KIM 2005). Mitotic progeny of 70-15 were tested to determine if telomeres were relatively unstable in comparison with other regions of the chromosome, and whether the mitotic progeny of 70-15 were more stable than the mitotic progeny of LPKY97-1A.

To determine if mitotic progeny had more restriction fragment length polymorphisms at telomeres than in other areas of the genome, the DNA from the third generation of single spores from LpKY97-1A was digested with *Eco*RI and

analyzed by Southern analyses subsequently with a telomere probe and then a probe from the Pot2 transposon (Figure 4-3). There is restriction site for *Eco*RI within MoTeR1 element, but not in MoTeR2. If MoTeR1s were located in the terminal position (i.e. the nearest element to the end of the chromosome) at multiple chromosome ends an intensely hybridizing telomeric band would be expected at approximately 1.4 kb in the Southern analysis.

Telomeres were significantly more variable than Pot2 loci in the third generation mitotic progeny of LpKY97-1A (p-value <0.001). There were 3.3 times more band changes observed in the telomere profiles than was observed with the Pot2 probe (Table 4-3). After three spore generations, the rice-infecting isolate 70-15 showed no band changes with the Pot2 probe and a single weakly hybridizing band was absent in 16/19 mitotic progeny when the Southern blot was probed with the telomere probe (marked with an asterisk in Figure 4-3).

In general more variability in telomere restriction fragments was seen in the Southern analyses of the *Eco*RI digests (Figure 4-3) than in the previous *Pst*I digests (Figure 4-1-D). In LpKY97-1A, a higher number of band changes were detected in the Southern analyses of *Eco*RI digests with 124 telomere band changes observed, while there were only 73 telomere band changes observed in the Southern analyses of *Pst*I digests. This suggested that telomere instability was underestimated in the previous Southern analysis of *Pst*I fragments.

4.3 Discussion

4.3.1 Instability of telomere restriction fragments in mitotic progeny of GLS isolates

One of the objectives of the experiments in this chapter was to determine if there was a difference in the stability of telomere restriction fragments in mitotic progeny between GLS isolates and the rice-infecting strain 70-15. LpKY97-1A generated new telomere fingerprints rapidly, while the rice-infecting strain 70-15 had very little change in the telomere fingerprints of mitotic progeny. The low telomere variability in progeny of 70-15 is similar to results that were based on RFLP profiles of subtelomeric elements (PARK *et al.* 2010).

Based on Southern analysis, new telomere fingerprints arose in mitotic progeny of the GLS isolates KS320 and RGNJ, which was also observed in LpKY97-1A. This suggested that the high variability of the telomeres previously observed from GLS isolates (FARMAN and KIM 2005) is based on underlying genetic instability at the chromosome ends. There also appears to be varying degrees of telomere instability between different GLS isolates because the frequency of band changes was higher in RGNJ and LpKY97-1A than in KS320 (Table 4-2).

Variability of restriction fragments has been observed in telomere-adjacent sequences from other organisms such as *Saccharomyces cerevisiae* (LOUIS and HABER 1990), *Plasmodium falciparum* (CORCORAN *et al.* 1988), *Lycopersicon esculentum* (BROUN *et al.* 1992), and *Xenopus laevis* (BASSHAM *et al.* 1998). Repeated sequences near the chromosome ends are thought to play a role in the observed variability of restriction fragments near the chromosome ends in these organisms. GLS isolates

harbor MoTeRs within their telomeres, and they likely drive telomere instability. MoTeRs are not only retrotransposons that could actively transpose to a new end, but the structure of these elements may lead to further instability. MoTeRs hybridized with all of the telomeric fragment gains in RGNJ and one of the two gains in KS320 (data not shown). The differences in relative stability between the isolates could be due to varying lengths of short internal telomere tracts between MoTeRs. These interstitial telomeres between the MoTeRs could cause stalling of the replication fork during DNA replication leading to double strand breaks (DSB) similar to what occurs at interstitial telomere tracts in *S. cerevisiae* (IVESSA *et al.* 2002). Long stretches of telomere (CCCTAA) or telomere variant repeats (CCCAAA and CCCGAA) are found at the 5' end of the MoTeR. These sequences could lead to increased instability when located between elements. A few of the RGNJ MoTeR-MoTeR junctions had long stretches of telomere repeats, while junctions between MoTeRs in LpKY97-1A had longer stretches of the variant telomere repeat (CCCAAA) (Table 3-4). Presumably these could represent potential double strand break points caused by inefficient DNA replication leading to instability of these telomeric fingerprints. Truncations of the telomere are known to occur frequently in *M. oryz*ae (ORBACH *et al.* 2000).

Another mechanism of increased variability could be caused by ectopic recombination between MoTeRs arrayed within telomeres. For example, in *Saccharomyces cerevisiae,* restriction fragments associated with the subtelomeric Y' elements undergo frequent change (HOROWITZ and HABER 1985; HOROWITZ *et al.* 1984). The restriction fragments changed due to mechanisms that included a

duplication of a Y' element by ectopic recombination between Y' at different chromosome ends and an unequal sister chromatid exchange event. Ectopic duplications were able to add a Y' on ends that previously did not have a Y' repeat. Tandem duplications of the Y' elements were seen to expand or contract. Gene conversion was also involved in the loss of restriction fragments (LOUIS and HABER 1990). *M. oryzae* isolates with longer MoTeR arrays or higher copy numbers of MoTeRs might be more prone to ectopic recombination. RGNJ and LpKY97-1A have higher MoTeR1 copy numbers than KS320 (data not shown), further supporting this argument.

Homologous recombination between subtelomeric repeats has been shown to cause contraction and expansion in repeat arrays (CORCORAN *et al.* 1988). The possible expansion of MoTeRs arrays by homologous recombination would lead to more interstitial telomere repeats, and could thus lead to further instability in the isolates.

Alternatively, higher instability in these isolates could be due to a deficiency in protein involved in telomere maintenance. For instance, deficiency in the yKU protein in yeast leads to an increase in homologous recombination between the subtelomeric Y' elements (FELLERHOFF *et al.* 2000; MARVIN *et al.* 2009). However, based on segregation analysis of a cross between a GLS isolate, and a lab isolate this in not likely (STARNES *et al.* 2012).

The impact of this high telomeric instability on the GLS isolates is not known. No aberrant colony morphologies or loss of aggressiveness was observed in mitotic progeny of LpKY97-1A in the laboratory (data not shown). The GLS pathogens are

highly aggressive on perennial ryegrass in the field (VINCELLI 1999), even with the high levels of telomere instability previously observed (FARMAN and KIM 2005).

4.3.2 Comparative instability of telomere and Pot2 restriction fragments

Another experimental objective was to compare the relative stability in telomere and Pot2 fingerprints in the mitotic progeny of 70-15 and LpKY97-1A. Both the telomeric and Pot2 fingerprints displayed more instability in the GLS isolate LpKY97-1A than was detected in 70-15. It was conceivable that telomere fingerprints would be unstable in GLS isolates, but the difference in instability between 70-15 and LpKY97-1A in the Pot2 fingerprints was unexpected. One assumption might be that the chromosome undergoes mutations at a constant rate regardless of position. The difference in stability between the two markers could be due to the greater likelihood of survivorship among progeny with telomeric mutations versus progeny with internal mutational events (RICCHETTI *et al.* 2003). Germination rates of the spores were not calculated, and only germinated spores were collected. This may suppress the ability to observe internal chromosomal instability due primarily to deleterious mutations at internal regions causing the spores to fail to germinate. Alternatively, the Pot2 polymorphisms could be due to Pot2 elements being present in chromosomal regions that are unstable in LpKY97-1A. Transposons tend to be found in clusters within the genome of *M. oryzae* (NITTA *et al.* 1997; THON *et al.* 2004) and these clusters frequently show multiple insertions from different transposable elements (FARMAN *et al.* 1996b; KACHROO *et al.* 1994; KACHROO *et al.* 1995). Transposable element clusters tend to have higher rates of gene duplications and gene evolution (THON *et al.* 2006).

However one other important factor must be taken into account in the presumed instability of Pot2 in LpKY97-1A, most of the Pot2 band gains observed in the Southern blot had weak hybridization signal. These are typically indicative of incipient Pot2 fragments that were developing during vegetative culture. Removing these band gains from the analysis led to no statistical support for a difference between Pot2 instability between LpKY97-1A and 70-15. The strong hybridization signals in the Southern hybridization with Pot2 did not change in LpKY97-1A after three generations (Figure 4-3). This would suggest that the majority of the genome is relatively stable. The extreme variability observed in telomeres may not correlate with the stability of internal loci or subtelomeres in the GLS isolates, which has been observed previously (FARMAN and KIM 2005). To test whether subterminal regions are unstable in LpKY97-1A, more subterminal probes are needed for use in Southern analysis.

MoTeR1 (~5 kb) and MoTeR2 (~1.7 kb) are substantially longer than the telomere repeats $(CCCTAA)_n$. The expansion and contraction of the telomere repeats may not be apparent in Southern analysis with *Pst*I or *Eco*RI digested DNA, but the changes in MoTeRs within these telomeres would be easily visible. Normal telomere turnover would be much more apparent in the GLS isolates than in 70-15 which lacks MoTeRs in the telomeres. Further work could address telomere turnover rate at the terminal telomere repeats of LpKY97-1A and 70-15.

Table 4-1. Telomeric *Pst*I restriction fragment changes of three generations of mitotic progeny in LpKY97-1A

Mitotic Progeny (G1)	Band Gains	Band Losses	Mitotic Progeny (G2)	Band Gains	Band Losses	Mitotic Progeny (G3)	Band Gains	Band Losses
1	0	0	1	1	1	1	1	0
2	0	0	2	5	4	2	2	0
3	0	0	3	2	2	3	1	0
4	1	0	4	1	0	4	6	4
5	1	0	5	2	2	5	4	2
6	1	0	6	2	0	6	4	2
7	1	0	7	0	1	7	2	1
8	0	1	8	4	1	8	4	2
9	0	0	9	1	0	9	3	2
10	1	0	10	3	2	10	1	0
11	0	0	11	0	0	11	3	3
12	0	0	12	1	0	12	4	2
13	0	0	13	0	2	13	1	0
14	0	0	14	1	1	14	2	0
15	0	0	15	1	1	15	2	0
16	1	0	16	5	4	16	3	3
17	0	0	17	4	1	17	3	3
18	0	0	18	1	1	18	2	0
19	0	0	19	2	3	19	2	1
Total	6	1	Total	36	26	Total	50	25
Average[1]	0.32	0.05	Average[1]	1.89[(a)]	1.37[(a)]	Average[1]	2.6[(a)]	1.32[(a)]

A superscript (a) indicates a significant difference at a p-value <0.001 from G1. Superscript 1 means that the average number of changes calculated per isolate.

Table 4-2. Telomeric *Pst*I restriction fragment changes in three GLS isolates and one rice infecting isolate.

Mitotic	RGNJ		KS320		LpKY97-1A		70-15	
Progeny	Gains	Losses	Gains	Losses	Gains	Losses	Gains	Losses
1	0	0	0	0	0	0	0	0
2	0	0	1	1	0	0	0	0
3	0	0	0	0	0	0	0	0
4	1	0	0	0	1	0	0	0
5	0	0	0	0	1	0	0	0
6	0	0	0	0	1	0	0	0
7	0	0	0	0	1	0	0	0
8	0	0	0	0	0	1	0	0
9	1	0	0	0	0	0	0	0
10	0	0	1	1	1	0	0	0
11	0	0	0	0	0	0	0	0
12	1	1	0	0	0	0	0	0
13	1	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0
15	1	0	0	0	0	0	0	0
16	0	0	0	0	1	0	0	0
17	0	0	0	1	0	0	0	0
18	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0
Total	5	1	2	3	6	1	0	0
Average	0.26[a]	0.05	0.11	0.16	0.32[b]	0.05	0	0

Superscript (a) indicates a significant difference from 70-15 at a p-value<0.05. Superscript (b) indicates a significant difference from 70-15 at a p-value<0.01.

Table 4-3. Changes in *EcoR*I restriction fragments with Pot2 probes and telomere probes

	LpKY97-1A					70-15			
	Pot2		Telomere			Pot2		Telomere	
Sample	Gains	Losses	Gains	Losses	Sample	Gains	Losses	Gains	Losses
G0	0	0	0	0	G0	0	0	0	0
LP3-1	0	2	0	0	R3-1	0	0	0	1
LP3-2	0	2	0	1	R3-2	0	0	0	0
LP3-3	0	1	0	1	R3-3	0	0	0	1
LP3-4	0	3	7	5	R3-4	0	0	0	1
LP3-5	0	3	6	7	R3-5	0	0	0	0
LP3-6	0	2	7	5	R3-6	0	0	0	1
LP3-7	0	1	1	2	R3-7	0	0	0	0
LP3-8	1	2	6	5	R3-8	0	0	0	1
LP3-9	1	2	6	6	R3-9	0	0	0	1
LP3-10	1	1	1	2	R3-10	0	0	0	1
LP3-11	1	1	5	4	R3-11	0	0	0	1
LP3-12	1	1	6	5	R3-12	0	0	0	1
LP3-13	0	1	5	4	R3-13	0	0	0	1
LP3-14	1	1	1	0	R3-14	0	0	0	1
LP3-15	1	1	1	0	R3-15	0	0	0	1
LP3-16	1	1	4	5	R3-16	0	0	0	1
LP3-17	1	2	5	8	R3-17	0	0	0	1
LP3-18	0	1	1	3	R3-18	0	0	0	1
LP3-19	1	1	2	1	R3-19	0	0	0	1
TOTAL	10	29	64	64	Total	0	0	0	16
AVERAGE	0.53[a]	1.53[a]	3.37[ab]	3.37[ab]	AVERAGE	0.00	0.00	0.00	0.84

Superscript (a) indicates a significant difference at the p-value<0.001 between LpKY97-1A and 70-15. Superscript (b) indicates a significant difference at the p-value <0.001 between telomere and Pot2 band changes.

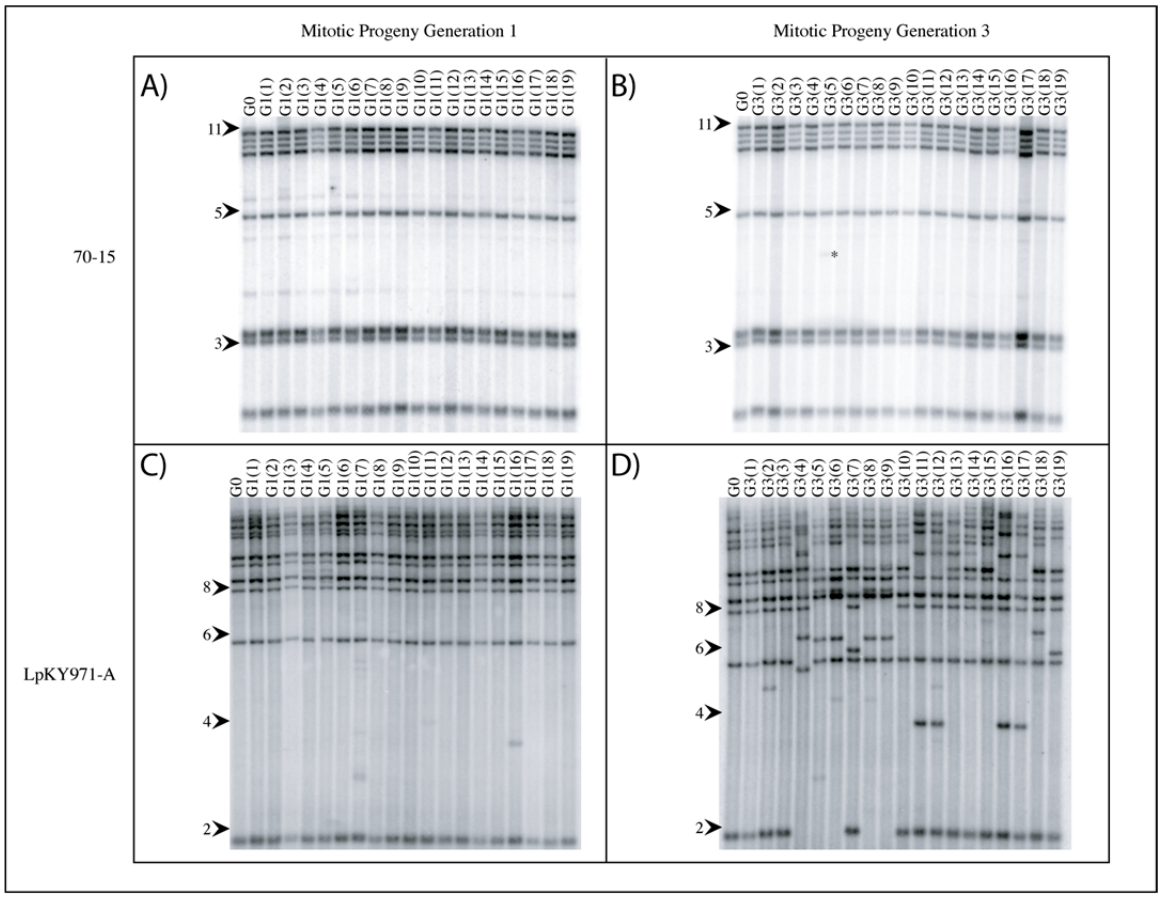

Figure 4-1. Southern analyses comparing stability of telomeric DNA fingerprints in 70-15 and LpKY97-1A. Genomic DNA was digested with *Pst*I, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized with a telomere probe. Black arrows indicate the size markers in kilobases. Lane 1 is the initial starting culture DNA fingerprint (G0). The first generation of spores (G1) from culture in 70-15 (A), and LpKY97-1A (C) are represented in lanes 2-20. The third generation of spores from plants in 70-15 (B) and LpKY97-1A (D) are represented in lanes 2-20. The asterisk represents a single faint *de novo* telomere band in 70-15.

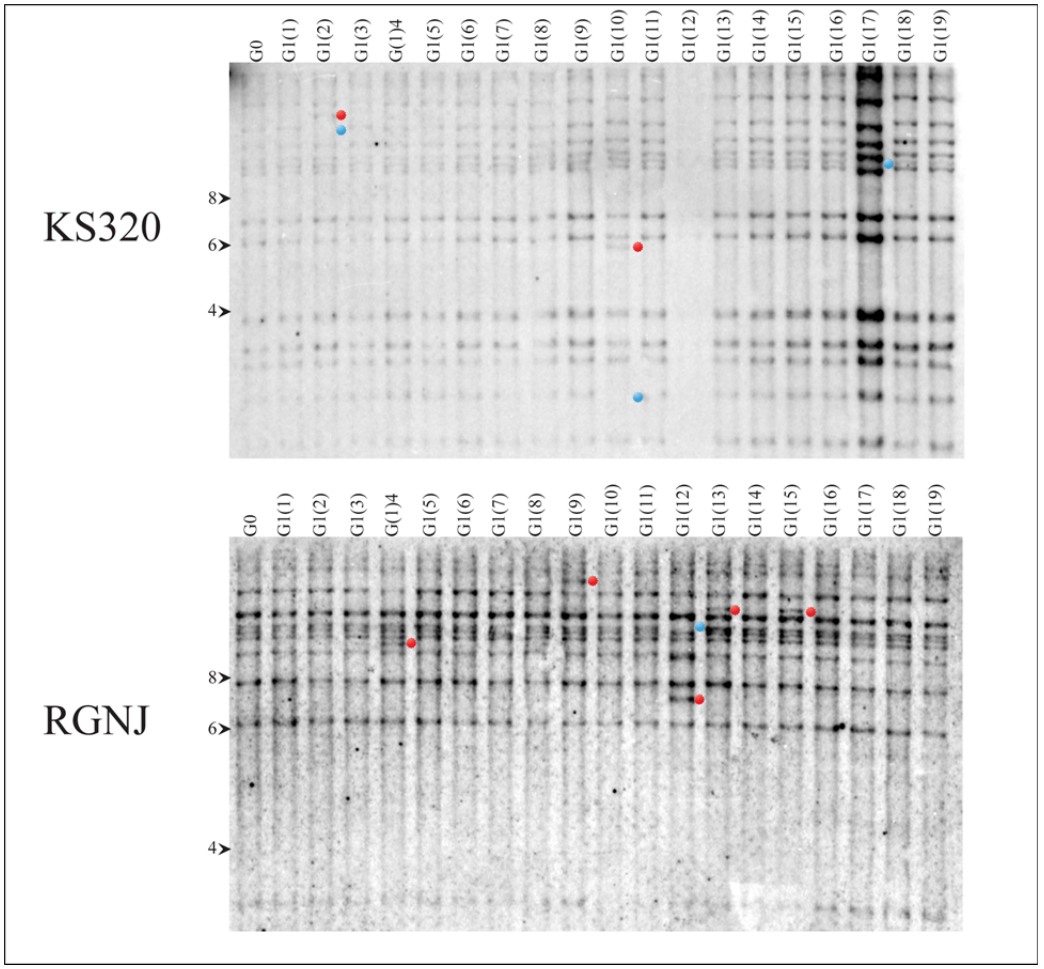

Figure 4-2. Telomere fingerprint stability of mitotic progeny from KS320 and RGNJ, which are Lolium pathotype isolates of *M. oryzae*. Genomic DNA was digested with *Pst*I, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized with a telomere probe. Black arrows indicate the size markers in kilobases. Lane G0 is the DNA fingerprint of the starting culture. Lanes 1-19 represent the DNA fingerprints of 19 spores collected from the initial culture. Red asterisks indicate a telomeric band gain, and blue asterisks indicate a telomere band that was lost.

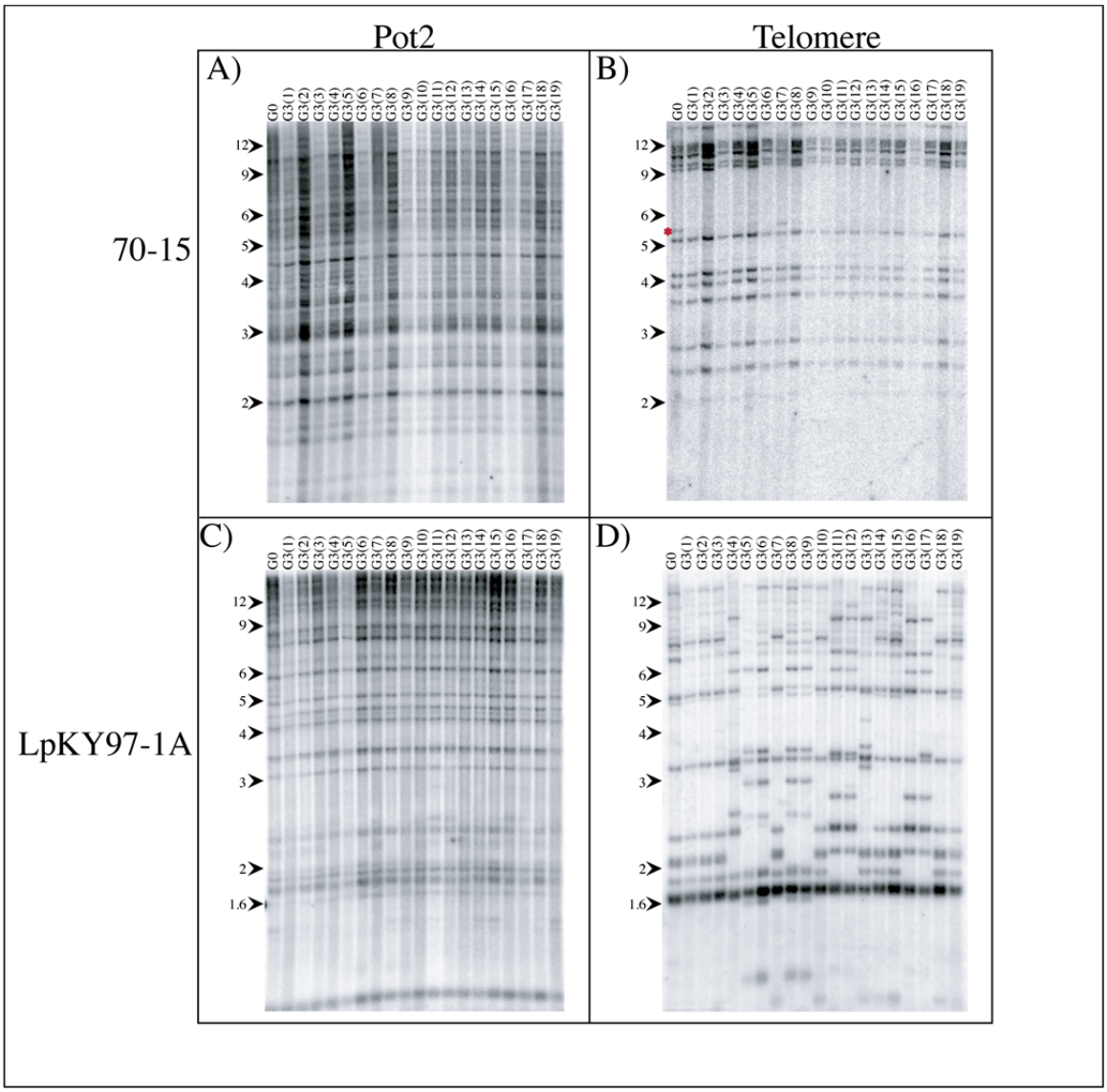

Figure 4-3. Southern analyses of Pot2 probe and telomeric probes in 70-15 and LpKY97-1A. Genomic DNA was digested with *Eco*RI, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized sequentially with Pot2 (A and C) and telomere (B and D). Between hybridizations blots were stripped. Black arrows to the left of the image indicate the size in kilobases. Lane one indicates the initial starting culture (G0). Lanes 2-20 represent third generation spores (G3) of 70-15 collected from the 51583 *Oryza sativa* cultivar (A and B). Lanes 2-20 in C and D represent third generation spores (G3) collected from the Linn variety of *Lolium perenne*. The red asterisk in lane 1 of Figure B represents a telomeric band that was lost frequently in mitotic progeny.

CHAPTER FIVE

Molecular Basis for Telomere Variability in LpKY97-1A

5.1 Introduction

Extreme variability is seen in telomere fingerprints between Lolium pathotype isolates of *Magnaporthe oryzae* (FARMAN and KIM 2005). The experiments from Chapter 4 revealed that the variability is readily observed among mitotic progeny of this pathotype. The Lolium pathotype isolate LpKY97-1A generated new telomere restriction fragment (trf) profiles rapidly, while the laboratory strain 70-15 underwent very little change in trf profiles after three spore generations (Figure 4-1). Similarly, other Lolium pathotype isolates, KS320 and RGNJ, generated new trf profiles in culture (Figure 4-2). The molecular basis for the telomere instability was not known.

MoTeRs are found in the telomeres of the Lolium pathotype isolates, but were not detected in 70-15 (Figure 3-7). It was believed that the activity of MoTeRs could account for the trf profile variation. Telomere fingerprint variation has been observed among mitotic progeny in *Saccharomyces cerevisiae*. The variation was due to telomere length heterogeneity and recombination among subtelomeric repeats (HOROWITZ *et al.* 1984; SHAMPAY and BLACKBURN 1988). The diversity of surface antigens in Trypanosomes is primarily the result of subtelomeric gene conversion and the frequency of variant surface glycoprotein switching is affected by telomere length (HOVEL-MINER *et al.* 2012). The experiments described in this chapter sought to address possible molecular mechanisms that are causing the trf

polymorphisms observed in the mitotic progeny of LPKY97-1A and how MoTeRs might be impacting the telomere stability.

5.2 Results

5.2.1 MoTeRs association with *de novo* telomere profiles in LpKY97-1A

In third generation progeny of LPKY97-1A *de novo* telomere profiles were observed in Southern analysis (Figure 4-1-D). It was unknown whether MoTeRs were in the telomeres of the *de novo* bands observed in the Southern analysis of mitotic progeny because some fragments were smaller than would be expected if a full length MoTeR1 was present. To determine if MoTeR1 or MoTeR2 were present in the *de novo* telomere fragments of mitotic progeny, DNA from 19 progeny (G3(1-19)) and the LpKY97-1A starting culture (G0) was analyzed by Southern analysis. The DNA was digested with *Pst*I, electrophoresed on a 0.7% agarose gel, and transferred to a nylon membrane. The Southern Blot was then probed with a telomere-specific, a MoTeR1-specfic, and a MoTeR2-specific probe. The blot was stripped between hybridizations with the different probes. Results are shown in Figure 5-1. The telomere fingerprint polymorphisms of the mitotic progeny were tabulated. A new telomere band in the DNA fingerprint of mitotic progeny that was absent in the parental isolate was counted as a gain, while the absence of a band in the DNA fingerprint of mitotic progeny was counted as a loss. *De novo* telomere fragments in mitotic progeny that cohybridized to MoTeRs were described as MoTeR associated gains (MAG). Telomere fragments that were absent in mitotic progeny, but that had previously cohybridized to MoTeRs in the starting culture were described as MoTeR associated losses (MAL). Results are summarized in Table 5-1.

Most of the new telomere fragments observed in the mitotic progeny (32 out of 50) had MoTeRs in their telomere based on cohybridization in the Southern analysis (Table 5-1). No telomere fragment below 4 kb cohybridized with the MoTeR probes (Figure 5-1). There were 25 telomere band losses in the Southern analysis of mitotic progeny and a majority of them (20 out of 25) were MAL. The five bands that were not MAL correspond to losses of the rDNA telomere fragment (Figure 5-1-A band R).

There was a non-telomeric *Pst*I fragment that hybridized to both MoTeR1 and MoTeR2 in G0 and in 15 of the 19 mitotic progeny (indicated by red arrows in Figure 5-1-B and 5-1-C). In the DNA fingerprint of progeny isolate G1(9) a gain of a non-telomeric MoTeR1 was observed (marked by an asterisk, Figure 5-1-B).

5.2.2 Methylation in telomere profile changes

*Pst*I, which recognizes the sequence CTGCAG, is sensitive to cytosine methylation, ^{5m}CTGCAG and CTGC^{6m}AG (McClelland *et al.* 1994). Therefore I speculated that methylation of genomic DNA could play a role in telomere fragment length polymorphisms observed in the Southern analyses. To determine if this was the case, two isoschizomers were used for comparative digests of genomic DNA prepared for Southern analyses. The restriction enzyme *Kpn*2I(*Bsp*MII) is sensitive to cytosine methylation and will not cut at the methylated sites T^{m5}CCGGA and TC^{m5}CGGA, while *Acc*III is insensitive to cytosine methylation (Labbé *et al.* 1988). For the purposes of this experiment three different LpKY97-1A strains (1G0, 1G2(5), and 1G3(4)), which had shown telomere length polymorphisms, were activated on OA. The strains were then treated as described in Section 2.4 for large scale DNA

preparation. The DNA was digested with *Pst*I, *Acc*III, or *Kpn*2I. The digested fragments were separated by electrophoresis on a 0.7% agarose gel, electroblotted to a nylon membrane, and probed with telomere probe.

1G0, 1G2(5), and 1G3(4) had several telomere restriction polymorphisms in the Southern analysis of *Pst*I digests. 1G2(5), as seen in lane 2 Figure 5-2-A, had three visible telomeric band changes from the initial starting culture. 1G3(4) exhibited at least nine telomeric band changes from the starting culture in the Southern analyses of *Pst*I digests. These data show that there are telomere restriction fragment polymorphisms in *Pst*I digests between the three strains.

Comparative analyses of the DNA fingerprints with restriction enzymes that were differentially methylation sensitive yielded no visible band changes in the LpKY97-1A isolates (Figure 5-2-B lanes three to eight). For example, 1G0 had 16 visible telomere bands in Southern blots of the *Acc*III/*Kpn*2I digests, with no telomere restriction fragment length polymorphisms between the two enzyme digests (Figure 5-2-B, lanes 4 and 5). There was DNA fingerprint variation between different strains using the methylation insensitive restriction enzyme (*Acc*III). For instance, the telomere fingerprints of 1G0 with 1G2(5) showed five distinct band differences. 1G3(4) had a greater number of band changes from 1G0 with a total of seven band changes observed.

5.2.3 Characterization of *de novo* telomere fragments

In experiments from Chapter 4, it was observed that the telomere restriction fragment profiles of GLS isolate LpKY97-1A were unstable, and *de novo* telomeres were readily observed in mitotic progeny (Figure 4-2). To determine some of the

possible mechanism involved in generating the telomere length polymorphisms a subset of strains from the previous studies were selected for further analysis. 1G0 is a single-spored culture of LpKY97-1A that was used as the starting strain for the experiments outlined in Chapter 4. Isolates 1G2(5) and 1G2(6) were collected from the second generation of mitotic progeny and 1G3(4) and 1G3(5) were collected from the third generation. When compared with one another all of these strains exhibited telomere restriction fragment polymorphisms. To describe the differences in more detail the DNA was digested with *Pst*I, electrophoresed on 0.7% agarose gel, electroblotted to a nylon membrane and probed with telomere and MoTeR1 probe. Blots were stripped between probings.

There are several examples of different *de novo* telomeres represented in Figure 5-3-A. In 1G2(5), a new telomeric DNA fragment was seen in Southern hybridization (Figure 5-3-A, lane 5, labeled with a B). This band did not cohybridize with a MoTeR1 (Figure 5-3-B, lane 5). The failure to cohybridize likely indicates that MoTeR1 was not responsible for the appearance of this new band. However, there were other examples in which MoTeRs did appear to play a role in the appearance of *de novo* telomeric fragments. In 1G3(4) and 1G3(5) new telomere fragments cohybridized with MoTeR1 (Figure 5-3, labeled with RM and BM respectively).

Shotgun cloning was used to capture a number of the telomeric restriction fragments. Genomic DNA was end-repaired and blunt end ligated into a plasmid vector. The ligation products were then digested with *Pst*I followed by ligation of the *Pst*I overhangs of the genomic DNA and plasmid vector. The resultant plasmids were electroporated into *E. coli*. Ampicillin was used to select for colonies that

contained plasmid inserts. Colonies that contained telomeres were identified by colony hybridization using a Telomere probe, and their plasmids were extracted. This process resulted in end-enriched plasmid libraries. The three large-scale DNA preparations 1G0, 1G2(5), and 1G3(4) were used to create three separate end-enriched plasmid libraries. The 1G0 end-enriched plasmid library contained the progenitor telomere restriction fragments; while the end-enriched plasmid libraries of 1G2(5) and 1G3(4) had the potential to yield a number of *de novo* fragments. This method yielded 51 total plasmids (9 from 1G0, 17 from 1G2(5), and 25 from G3(4)).

The inserts of all plasmids were partially sequenced to determine if the telomere tracts were at either end of the insert. Though all inserts had telomere sequence, not all had the long telomere repeat tracts associated with chromosome ends. In the end-enriched library of 1G0, 5 out of the 9 plasmids had telomere tracts. Only 8 out of the 17 inserts in 1G2(5) plasmid libraries had telomere tracts. There were 20 out of 25 inserts, which contained telomere tracts in the end-enriched library of 1G3(4).

To determine the relatedness of the partially sequenced inserts to known elements and each other, the sequences obtained from the end-enriched plasmid inserts were analyzed by using local blast to search databases containing MoTeRs, the 70-15 genome, and the sequences of the other plasmid inserts. The rDNA chromosome ends were captured in all three of the plasmid libraries. Two chromosomes ends unique to LpKY97-1A were found in 1GL3(4) and were designated 31B and 117B1 (naming is based on the plasmid clone where the

sequence was obtained). The 31B chromosome end was captured in two plasmid clones. The 117B1 chromosome end was captured multiple times (16 of the 20 inserts in the 1G3(4) plasmid library).

Sequences that were found to be unique at a chromosome end were used to develop probes (31B001 and 117B1) for Southern analysis. This would provide the ability to track changes at specific chromosome ends in mitotic progeny. The data gathered from the Southern analysis of unique probes and sequencing were then used to characterize some of the possible mechanisms behind telomere instability in LpKY97-1A by following a few of the chromosome ends in 1G0, 1G2(5), and 1G3(4). The characterizations of the *de novo* telomeres are detailed below.

5.2.3.1 rDNA telomere truncations

The telomeric rDNA fragment (indicated with an arrow labeled with an R in Figure 5-1-A) was absent in five of the 19 third generation mitotic progeny. In the isolates used to prepare the end-enriched plasmid libraries both 1G0 and 1G2(5) have the same telomeric rDNA band, while this band was absent in 1G3(4) (Figure 5-2-A). The rDNA telomeric fragment, labeled R in Figure 5-1-A, was captured in the end-enriched plasmid library of both 1G0 and 1G2(5) and had been sequenced. Alignment of this sequence with rDNA sequence of 70-15 revealed that the rDNA was truncated at position 7579 and capped with telomere repeats, $(CCCTAA)_{29}$ (Figure 5-4-A). The rDNA telomere of 1G3(4) had been captured during the end enrichment process and both sides of the insert were partially sequenced. This revealed a 5' truncated MoTeR1 (position 2248) capped with a telomere $(CCCTAA)_{27}$ at the chromosome end and rDNA sequence at the subtelomere. PCR

was then used to characterize the MoTeR to rDNA junction. The junction between the rDNA and the 3' end of MoTeR1 was amplified by PCR and then sequenced. The sequence data revealed that MoTeR1 was located one CCCTAA repeat upstream of the same rDNA truncation described in 1G0 and 1G2(5) (Figure 5-4). This suggested that the MoTeR1 had inserted into the rDNA telomere. The rDNA fragment had likely shifted in 1G3(4), as observed in Southern hybridization in Figure 5-3-A, (lane 4 labeled RM). The MoTeR1 probed hybridized with the *de novo* telomere fragment (Figure 5-3-B, lane 4). The size increase was less than would be expected if an intact MoTeR1 had inserted into the rDNA telomere, and the band was similar in size to the rDNA telomere captured in 31B.

An alternative explanation to MoTeR1 insertion into the rDNA telomere, however, is that LpKY97-1A could have multiple chromosome ends with rDNA sequences, and that the MoTeR1-rDNA band (RM) observed in 1G3(4) could have resulted from the alteration of a MoTeR-containing rDNA telomere already present in the starting culture. To test this idea, PCR was used to test for the presence of a MoTeR1-rDNA junction in the starting culture. Using an rDNA and MoTeR 3' primer pair, DNA from 1G3(4) yielded a PCR amplicon of the expected size (data not shown), while DNA from 1G0 failed to produce any products. This indicated that there were no existing MoTeR-rDNA junctions in the initial starting culture.

5.2.3.2 Capture and duplication of internal sequence at a telomere

Southern analyses of *Pst*I digested DNA revealed that isolates 1G2(5) and 1G3(4) both have a *de novo* telomere band of 700 bp (Figure 5-3-A). Plasmid clone 117B1, from the end-enriched plasmid library of 1G3(4), had a similarly sized insert. The

plasmid insert was sequenced and local blast analysis against the genome sequence of 70-15 revealed no sequence homology to 70-15 other than to the telomere repeats. The telomere sequence found at the end of 117B1 band B was $(CCCTAA)_{29}(CCTAA)(CCCTAA)$. To determine if the plasmid insert in 117B1 was the telomere fragment labeled B in Figure 5-3-A, a probe was generated from the 117B1 sequence to be used for Southern analysis. Primers (m117B1F and m117B1R) were designed from the unique sequence of 117B1. PCR was then used to amplify the sequence from the plasmid 117B1 for use as a probe. The 117B1 probe was then hybridized to the previous Southern Blot (Figure 5-3-C). In 1G2(5) and 1G3(4), the telomeric fragment (Figure 5-3-A, labeled B) cohybridized with the 117B1 fragment (Figure 5-3-C) and thus these fragments represented the same *de novo* telomere captured in the plasmid 117B1. The initial culture (1G0) produced three strong hybridization signals with the 117B1 probe (Figure 5-3-C). None of the 117B1 signals coincided with the telomeric fragments, suggesting that they were internal genomic fragments. These internal fragments were also found in 1G2(5), 1G2(3), and 1G3(5) (Figure 5-3-3) indicating that these fragments were not lost in these isolates.

In 1G3(5) another strong hybridization signal was observed with the 117B1 probe (labeled BM, Figure 5-3-C, lane 5). The band in question also hybridized with MoTeR1 (labeled BM, Figure 5-3-B, lane 5) and was the approximate size expected if an intact MoTeR1 had transposed into the newly generated 117B1 chromosome end (labeled B, Figure 5-3-A lanes 2 and 4). None of the internal 117B1 bands appeared to be missing in this isolate (Figure 5-3-C lane 5) likely indicating that this fragment

did not arise from the truncation of another chromosome. To confirm that the novel BM telomere contained a MoTeR1, the junction between 117B1 and MoTeR1 was PCR-amplified from the genomic DNA of isolate 1G3(5) using the primers (117B1JF and 117B1JR). The resulting amplicon was sequenced, revealing a junction consisting of the telomeric motifs $(CCCTAA)_3$(CCTAA)(CCCTAA) between the 3' end of MoTeR1 and the 117B1 unique sequence (Figure 5-4).

5.2.3.3 Truncation of MoTeR2 arrays

The telomeric bands marked with an arrow (Figure 5-5-A lanes 2-20) showed variable hybridization intensities among isolates. This likely indicates rearrangements had taken place in the telomeric fragments that were associated with this band. To analyze the changes of the telomeres in mitotic progeny more closely, I needed to be able to track a single chromosome end. One plasmid in the end-enriched library of 1G(3)5, 31B, contained a 9.1 kb insert. Sequencing of the T3 end of the insert revealed the 5' end of MoTeR element capped with telomere repeats $(CCCTAA)_{30}$. The T7 end of the insert was sequenced and local blast analysis against the 70-15 genome failed to find significant matches indicating that the sequence was unique to LpKY97-1A. Primers (31B001F and 31B001R) were designed to PCR-amplify 600 bp of sequence from the T7 end of the insert. The resulting PCR amplicon, 31B001, was then used as a probe for Southern analysis of the telomere blot shown in Figure 5-5-A.

The 31B001 probe hybridized to two fragments (~9 kb and ~11 kb) in the starting culture (G0) (lane 1, Figure 5-5-B, labeled A and B respectively). The 31B1 telomere fragment obtained in the end-enriched library of 1G3(5) corresponds to

the weakly hybridizing B in the isolate G0, but the strong hybridization signal in G0 corresponds to Band A. The weak hybridization signal of Band B in G0 likely indicates that the culture was heterokaryotic for the 31B telomere, and that the shorter fragment had been cloned. In the 3rd generation of mitotic progeny, eight of the 19 progeny had the same 3100B1 fingerprint as G0 (Figure 5-5-B). Four isolates, G3(5), G3(6), G3(8), and G3(9), only had the ~9 kb variant of the 31B telomere. Isolate G3(4) had three bands in its 31B001 fingerprint (~13 kb, ~11 kb, and ~9 kb), with the ~13 kb band being the only strong hybridization signal. The regularity of the band size variants of ~2 kb suggested a genetic element was expanding/contracting at this chromosome end. The expected difference of sizes based on MoTeR2 element would be 1.7 kb, but the separation of the bands wasn't enough to obtain accurate sizes.

The structure of the 31B telomere restriction fragment had not yet been fully explored. Information about the structure of the MoTeRs in the fragment was needed to determine if MoTeR2 arrays were likely at this telomere. Restriction analysis and partial sequencing of the 31B plasmid insert revealed that downstream of the telomere repeats $(CCCTAA)_{30,}$ there was an intact MoTeR2 element separated from a highly 5' truncated MoTeR1 (position 4726) by a single CCCTAA motif. The truncated MoTeR1 element was itself separated from the unique 31B chromosome end by a single CCCTAA repeat (Figure 5-4).

An *Apa*I restriction site was present at ~1.9 kb upstream of the truncated MoTeR1. It was believed that by digesting the genomic DNA of the mitotic progeny with *Apa*I, prior to Southern blotting, a better resolution in the sizes of the variant

31B telomere fragments could be obtained. Therefore, primers (31B002F and 31B1002R) were designed for PCR amplification of an 868 bp probe fragment (31B002) from the unique DNA upstream of the truncated MoTeR1, since the original probe no longer hybridized with 31B telomeric fragments in *Apa*I digests. A representative set of DNA from isolates having different 31B fingerprints were digested with *Pst*I or *Apa*I prior to Southern blotting and then probed with 3100B2. The resulting Southern blots were then compared.

The 31B002 DNA fingerprint profiles (Figure 5-6, lanes 10-18) in the *Pst*I digested samples were the same as previously observed for those isolates (shown in Figure 5-5-B), except for 2 weak hybridization signals. In G3(6) the weak hybridization signal was at ~7.4 kb, which is below the 9.1 kb fragment observed in the previous blot probed with 31B001 (lane 12, Figure 5-6). This is exactly the size change that would be expected if the 9.1 kb fragment had lost a MoTeR2 in the 31B telomere. However, this band was not present in the Southern analysis of *Apa*I digested DNA (lane 3, Figure 5-6) likely indicating a chromosomal rearrangement proximal to the *Apa*I site. In isolate G3(13), a large weak signal was seen in the 31B002 Southern hybridization that was absent in the previous blot probed with 31B001.

*Apa*I digestion provided a better resolution of the 31B band changes (lanes 1-9, Figure 5-6). Several new weakly hybridizing bands were also observed (lanes 1-9, Figure 5-6). The longer 31B002 probe appears to be more sensitive than the 31B001 probe.

In G0 the most intensely hybridizing fragment (A) was ~5.8 kb in length and the weakly hybridizing fragment (B) was ~4.1 kb (Figure 5-6). Since MoTeR2 is ~1.7 kb in length, the size difference between bands A and B was exactly what would be expected if the 31B002 fragment contained an extra MoTeR2 copy. Another weakly hybridizing fragment was observed in G0, which was ~ 5 kb in length (Figure 5-6-1, lane 1, labeled C). This fragment was not seen in the previous Southern blot of *Pst*I digests probed with 31B002, and could suggest the loss the *Apa*I site 1.9 kb from the start of the tMoTeR1.

In G3(6), only band B, which is the shortest MoTeR2 array variant, was observed in the Southern analyses of *Apa*I digested genomic DNA with the probe 31B002 (Figure 5-6, lane 3). G3(13) had a 31B002 profile with four bands, each of these were at intervals of ~1.7 from the longest fragment at ~9.2 kb (~4 MoTeR2s in the array) to the smallest at ~4.1 kb (Band A). The dominate 31B fragment in this isolate was ~7.5 kb in size, which likely represents ~3 MoTeR2s arrayed in this telomere.

It was believed that the longer MoTeR2 arrays would continue to truncate down to the single MoTeR2 in the telomere (represented by Band B, Figure 3-6), which would remain stable. To test this idea, 7 different third generation progeny with variant 31B telomeres were grown from paper disks placed on the center of separate oatmeal agar plates. The isolates were allowed to grow across the plates, and spores were harvested. A total of 12 germinated spores from each isolate were cultured, and then DNA was extracted from at least 10 progeny per isolate. The DNA samples were digested with *Apa*I, the restriction fragments were separated by

agarose gel electrophoresis, and electroblotted to a nylon membrane. The blots were stripped between each probing and the results are shown in Figures 5-7 and 5-8.

G3(6) had a strongly hybridizing fragment at ~4.1 kb (Band B), with an array of a truncated MoTeR1 and a MoTeR2 in the 31B telomere. The 31B002 probe revealed no changes in the 12 progeny surveyed (Figure 5-7, lanes 3-14). Telomere fragment variation was observed in these isolates (Figure 5-7-B), but the 31B telomere was stable.

G3(14) and G3(15) had a presumed array of the truncated MoTeR1 followed by ~2.01 MoTeR2s in the 31B telomere (Novikova et al. unpublished), which was ~5.8 kb in length (Band A). In 3 of the 21 (or 14%) mitotic progeny of G3(14) and G3(15), Band B was the predominant variant, while in the other 18 progeny a weakly hybridizing Band B was seen along with the strong hybridization signal for Band A (Figure 5-8).

G3(11), G3(12), G3(13), and G3(16) had intense signals at ~7.5 kb, which was believed to have a truncated MoTeR1 and ~3.01 MoTeR2s in the 31B telomere (Novikova et. al. unpublished). Six of the 44 progeny (or 13%) had intense signals at sizes different from their respective starting strain (Figure 5-7). G3(12-2) and G3(13-2) had the A variant representing a truncation of one MoTeR2, while G3(12-12) had a strongly hybridizing fragment at ~4.1 kb or a loss of two MoTeR2 elements. One progeny, G3(11-4), had a truncation of an unexpected size of ~6.7 kb. G3(11-5) and G3(11-6) had longer fragments (~9.2 kb and ~15 kb respectively) likely representing further extensions of the MoTeR array in the 31B telomere.

5. 3 Discussion

5.3.1 MoTeRs in telomere restriction fragment length polymorphisms

The major objective of the experiments in this chapter was to characterize the possible mechanisms by which telomere restriction fragment length polymorphisms could arise in LpKY97-1A. Eight of the nine visible telomere fragments in *Pst*I-digested genomic DNA cohybridized with MoTeRs, and they were associated with a majority of gains (64%) and losses (80%) of telomeric bands in mitotic progeny. This close association of MoTeRs with many of the telomere restriction fragment length changes in mitotic progeny suggested that MoTeRs had a role in this instability. However, a direct correlation with band changes and MoTeRs activity could not be established due to several possible factors such as: telomere fragments of the same size comigrating during agarose gel electrophoresis and appearing as a single band in the Southern analysis, differential methylation, MoTeR truncations caused by other mechanisms, or another transposable element causing the instability of chromosome ends. The cloning of telomere restriction fragments was necessary to establish a direct link of band changes with MoTeRs. Shotgun cloning led to the creation of end-enriched plasmid libraries from three different LpKY97-1A strains. The chromosome ends captured by this approach were used to characterize three different mechanisms by which instability of telomere fingerprints can arise: MoTeR transposition into telomeres, duplication and capture of internal sequence at a telomere, and expansion or contractions of MoTeRs arrays.

However, there were other possible mechanisms that needed to be explored, as MoTeRs cohybridization was absent in 36% of the telomere restriction fragment

length gains. Methylation has been known to play a role in restriction fragment length polymorphism at chromosome ends (LANGE *et al.* 1990). The effect of methylation was ascertained by comparing telomere fingerprints of genomic DNA samples digested with restriction enzyme isoschizomers, one being sensitive to cytosine methylation and the other being insensitive to methylation. The telomere fingerprints failed to yield any band differences. Thus, cytosine methylation may not be an important mechanism in telomere restriction length polymorphisms in LpKY97-1A, which suggests that other mechanisms could account for the band changes.

There is some evidence to suggest that other retrotransposons are active. The retrotransposon MAGGY, for example, has been shown to be active and has been linked to some telomere profile changes in the mitotic progeny of LpKY97-1A (NOVIKOVA *et al.* 2011). Thus, other elements could be actively causing telomere instability within the GLS isolates. The structure of MoTeRs may serve as a hotbed for other transposable elements, which could lead to further instability of the telomeres.

5.3.2 rDNA rearrangements

The telomeric rDNA band in Southern hybridizations (labeled R in Figure 5 1-A) was lost in 14 of the 58 mitotic LpKY97-1A progeny (24%) over the course of three generations (data not shown), indicating that the telomeric region of the rDNA array is unstable in this isolate of *M. oryzae*. Instability in the telomeres at the end of the rDNA arrays has also been seen in the related filamentous fungus *Neurospora crassa* (WU *et al.* 2009).

In LpKY97-1A the instability of a characterized rDNA band was potentially caused by the activity of a MoTeR1. In the mitotic progeny, 1GL3(4), the absence of the telomeric rDNA band (Figure 5-1-A lane 1 labeled R) was believed to have been produced by a MoTeR1 that transposed into a telomere of the rDNA array. This caused the telomeric rDNA fragment to be longer (Figure 5-1-A lane 4 labeled RM). The size difference between the two bands was less than expected if an intact MoTeR1 had transposed into the rDNA telomere. Sequencing analysis revealed that a 5' truncated MoTeR1 (tMoTeR1) was in this telomere, and the truncation point of MoTeR1 was in register with the telomere repeats (CCCTAA)-tMoTeR. The tMoTeR1 could have become established in this telomere through retrotransposition or recombination.

There are several different models by which the tMoTeR1 could have been inserted into the rDNA telomere. The 5' end insertion mechanisms of non-long terminal repeat retrotransposons are not known. Truncation of 5' ends may occur by the reverse transcriptase (RT) failing to copy the entire RNA template or premature initiation of second-strand synthesis (LUAN *et al.* 1993). One current model suggests that annealing of microhomologies between the reverse transcribed cDNA to the top-strand will prime second-strand synthesis of the retrotransposon (STAGE and EICKBUSH 2009). If binding occurs within the cDNA instead of at the end, then a 5' truncated element would be inserted. A microhomology of one base pair was found between the MoTeR1 truncation point and the telomere sequence, but in order for the truncation to be in register with the telomere a 5 bp target site truncation (TST) or target site duplication (TSD) would have needed to occur.

Another possible scenario, is that the RNA template was derived previously from a tMoTeR1 from another telomere. This tMoTeR1 would potentially have a telomere repeat at the 5' end of its RNA transcript, could utilize the RT machinery of a full length MoTeR1, and the telomere repeats of the reverse transcribed cDNA would provide the needed microhomology for the top-strand insertion of the tMoTeR1. An alternative mechanism to explain the tMoTeR insertion is the template jump model (BIBILLO and EICKBUSH 2004; ZINGLER *et al.* 2005b). Based on this model, the MoTeR1 was incompletely reverse transcribed and the RT was able to template jump onto the top-strand to begin the second-strand synthesis reaction. One less likely mechanism is more complex. A full length MoTeR1 could have retrotransposed into the telomere repeat in one of the nuclei that gave rise to a mitotic progeny. During the further culturing of this isolate the MoTeR1 could have become truncated followed by addition of telomere repeats by telomerase. This mechanism is not likely because the integrity of terminal telomere repeats is typically maintained in *M. oryzae* (STARNES *et al.* 2012). The maintenance of telomere repeat caps has also been seen in another fungus, *Kluyveromyces lactis,* where all but the innermost telomere repeats undergo gradual turnover primarily through the loss of repeats by replicative processes and the addition of new repeats by the action of telomerase (MCEACHERN *et al.* 2002).

Recombination with another telomere end is an alternative mechanism to retrotransposition. This was considered to be an improbable mechanism whereby the tMoTeR1 was copied into the rDNA telomere. Analysis of the junction sequence between the rDNA array and the 3' end of tMoTeR1 revealed only one CCCTAA

repeat between the two, which is presumably too little homology for efficient homologous recombination (LISKAY *et al.* 1987; RUBNITZ and SUBRAMANI 1984). Break-induced replication (BIR), however, has been known to copy sequences to new chromosome ends even when there are only a few nucleotides of homology. This occurred at an extremely low frequency (RICCHETTI *et al.* 2003), which suggests that BIR is not the probable explanation for the insertion of the tMoTeR1 in the rDNA telomere. Furthermore, the justification that recombination was responsible for tMoTeR1 copy into the rDNA telomere would have required a significant degradation of the rDNA telomere tract or a double strand break (DSB) to initiate BIR. Thus recombination could account for the tMoTeR1 in the rDNA telomere, but it is much less likely than the transposition model.

5.3.3 Duplication and capture of internal sequence at telomeres

In two progeny of LpK97-1A there was a *de novo* telomere band (Lanes 2 and 4, labeled B, Figure 5-3-A). There are a number of ways this telomere could have been formed. One possible mechanism is that the chromosome end was truncated in these isolates, followed by the addition of telomere. However, this is not a probable mechanism for *de novo* telomere formation, because the variant isolates all retained the internal 117B1 bands (Figure 5-3-C). If a truncation of a chromosome end had occurred, it would be expected that one of the 117B1 internal bands seen in the original starting isolate would have been absent in that progeny. More likely, the new telomere arose through the copying of an internal fragment onto a truncated chromosome end. It is unknown what may have caused this truncation event, and the chromosome end has not been sequenced downstream of the *Pst*I site. A

truncation in a MoTeR array, whereby a short telomere sequence was left at the 3' end or an internal break within a MoTeR, may have initiated double strand break repair machinery, such as BIR, which could have been active in capturing the internal sequence. Further characterization of this chromosome end is needed to understand the potential mechanisms involved in the capture of internal sequence at the chromosome end.

5.3.4 MoTeR transposition

The telomeric 117B1 was not present in the initial isolate, but was present in at least two mitotic progeny. The difference between band B and band BM was approximately 5 kb, which is suggestive of an addition of a full length MoTeR1 into the 117B1 telomere (Figure 5-3-A). Sequencing revealed that at the junction between the 3' end of MoTeR1 and 117B1 in band BM was the sequence MoTeR1-$(CCCTAA)_3$(CCTAA)(CCCTAA)-117B1. The variant telomere sequence CCTAA was also found in the same position in band B. The 5' end of the MoTeR1 sequence and the telomere proximal sequence were not captured. Based on hybridization intensity of the telomere probe to fragment BM, it is expected to have a normal length telomere. These data suggested that the MoTeR1 was added into the telomere after the formation of the *de novo* 117B1 telomere. This could have been accomplished by retrotransposition of MoTeR1 or recombination of the telomere with a MoTeR1 containing telomere. Though as discussed earlier, recombination was not the favored mechanism.

There are several possible mechanisms by which the MoTeR1 could have retrotransposed into the 117B1 telomere. Our current model suggests that

telomere repeats are incorporated into the 3' end of the RNA transcript, which would allow for insertion of the MoTeR1 in register with the telomere repeats of the target site (STARNES *et al.* 2012). These telomere repeats may bind to the target site in the telomere and act as a primer to initiate reverse transcription of the MoTeR transcript by a process known as target primed reverse transcription (TPRT). The inclusion of target site DNA at the 3' end of the transcript is known to increase the accuracy of insertion, but can reduce the efficiency of TPRT (LUAN and EICKBUSH 1996). A variable number of telomere repeats are found at the 3' junctions of MoTeRs, which could be caused by the selection of different target sites in the telomeric repeats or a larger number of telomere repeats incorporated into the MoTeR transcript.

Preliminary evidence has shown that MoTeRs are expressed (appendix C). Expression analysis by reverse transcriptase PCR has utilized a MoTeR specific primer in the initial cDNA synthesis. Use of telomere repeats as a primer for cDNA synthesis in this assay could establish whether telomere repeats were part of the transcript. Cloning and sequencing of MoTeR1 reverse transcriptase PCR amplicons could provide evidence of read through transcription into the telomere repeats.

MoTeR1 is the first example of a potentially active telomere repeat specific non-LTR retrotransposon in fungi. Two examples have been provided that are highly suggestive of MoTeR retrotransposition. One example is the retrotransposition of a tMoTeR1 into the rDNA telomere and the other example is the retrotransposition of a suspected full length MoTeR1 into a *de novo* telomere.

However, further experiments are needed to confirm that MoTeR1 and MoTeR2 are capable of retrotransposition. The only other NLRs where direct evidence has shown that they actively insert into telomere repeats are the TRAS and SART retrotransposons from *Bombyx mori* (FUJIWARA *et al.* 2005).

5.3.5 Molecular mechanisms underlying MoTeR array contraction and extension

By following the individual chromosome end 31B in mitotic progeny of LpKY97-1A, it was possible to track the fate of MoTeR2 arrays. The smallest size 31B *Apa*I digested telomere fragment in third generation mitotic progeny of LpKY97-1A was ~4.1 kb in length (lane 3, Figure 5-5). Restriction analysis and partial sequencing of this fragment revealed that downstream of the terminal telomere repeats there was a MoTeR array with an intact MoTeR2 and a 5' truncated MoTeR1 in the telomere of 31B (Figure 5-4). The ~4.1 kb fragment was seen as a faint band in the Southern analysis of the *Apa*I-digested genomic DNA in the initial starting culture (labeled B, lane 1, Figure 5-6), which means it is likely only present in some nuclei. The *Apa*I-digested 31B chromosome end present in most nuclei within the original culture was ~5.8 kb. This is exactly the size difference expected if an additional MoTeR2 were present in the MoTeR array in the telomere of the 31B chromosome end. Third generation mitotic progeny with 31B fragments larger than ~5.8 kb were always accompanied with faintly hybridizing fragments of ~5.8 kb and ~4.1 kb (Figure 5-6, lanes 2,4-6, and 9), suggesting that truncations of the MoTeR2 array were common during mitotic division. Further Southern analyses of progeny from the isolates with 31B fragments larger than ~5.8 kb showed that truncations of the

MoTeR2 arrays occurred more frequently than expansions (Figure 5-7 and 5-8). Progeny of the isolate G3(6) with the ~4.1 kb fragment did not show any expansion or contraction of the MoTeR2 array in the Southern analysis of ApaI digests (Figure 5-7-A, lanes 3-15), demonstrating that this is a stable configuration of the telomere. It could be expected that this MoTeR2 array could truncate further by the loss of an additional MoTeR2. However, a further truncation of this array was not observed in any of G3(6) mitotic progeny.

One possible mechanism that could lead to truncation of MoTeR2 arrays is that the replication forks stall at long telomere tracts between MoTeR2 elements in an array, which then causes truncations in the MoTeR2 array. Truncations were more pronounced in the longer MoTeR2 arrays of the 31B telomere. Cloning and sequencing of the ~5.8 kb fragment revealed a long telomere tract [$(CCCTAA)_2(CCCTAAA)(CCTAA)_{12}$] between the two MoTeR2s in the array (Novikova *et. al.* unpublished). In *Saccharomyces cerevisiae* telomere-promoted replication fork stalling is dependent on the length of the telomeres (MAKOVETS *et al.* 2004). The instability observed could be the result of double strand breaks being created at stalled replication sites (SAINTIGNY *et al.* 2001; STRUMBERG *et al.* 2000). Chromosome instability is known to occur at interstitial telomere repeats (KILBURN *et al.* 2001; MUSIO *et al.* 1996), further suggesting that long interstitial telomere repeats between the MoTeRs in an array could lead to higher instability of telomere fingerprints.

The Interstitial telomere repeats in MoTeR may also facilitate the formation of telomere loops or "t-loops". T-loops are formed when a long stretch of double

stranded telomere DNA loops around and the single stranded terminus invades back into double stranded telomere forming a displacement loop (D-loop) (GRIFFITH *et al.* 1999). The topological barriers caused by a t-loop could lead to replication fork stalling (SAMPATHI and CHAI 2011). Processing of t-loops by homologous recombination in cells with mutant alleles of TRF2, a DNA binding protein involved in telomere protection, lead to truncations of the telomere at the site of the t-loops (WANG *et al.* 2004). Improper disassembly of t-loops not only leads to telomere truncations, but also to the formation of telomeric circles (VANNIER *et al.* 2012). These telomeric circles are known to stimulate recombinational telomere elongation (NATARAJAN *et al.* 2003; NATARAJAN and MCEACHERN 2002), which could lead to further instability in MoTeR containing isolates.

The presence of MoTeR sequence in the telomere may inhibit the proper function of telomere binding proteins leading to further instability in these telomeres. This could be an interesting avenue to explore with further research.

Another mechanism that explains both truncations and expansions of MoTeR2 arrays is unequal crossing over. This mechanism could lead to expansion in one MoTeR array and contraction of another MoTeR2 either in a non-homologous chromosome end or the sister chromatid. Expansion and contraction of the rDNA array in haploid cells can be caused by unequal crossing over of sister chromatids (SZOSTAK and WU 1980). Sister chromatid exchange occurs at a higher rate at telomeres and subtelomeres than in internal chromosomal sequence (RUDD *et al.* 2007). This further suggests that MoTeRs could be susceptible to high rates of recombination between sister chromatids. The MoTeR system could provide

additional evidence of the frequency by which sister chromatid exchange occurs. This could be accomplished by following a “marked” MoTeR element or following a particular variant of MoTeR through asexual reproduction cycles.

Table 5-1. MoTeR associated telomere restriction fragment changes in mitotic progeny of LpKY97-1A. Telomere restriction fragment changes were counted from Figure 5-1 (Gains and Losses). In the MoTeR Associated column, band changes that cohybridized with MoTeR were given a score of 1, and band changes that did not cohybridize were given a score of 0. In each column the number represents the total number of band changes for single isolate.

Sample	Gains	MoTeR Associated	Losses	MoTeR Associated
G3-1	1	1	0	0
G3-2	2	2	0	0
G3-3	1	1	0	0
G3-4	6	3	4	3
G3-5	4	2	2	1
G3-6	4	2	2	1
G3-7	2	2	1	1
G3-8	4	2	2	1
G3-9	3	2	2	1
G3-10	1	1	0	0
G3-11	3	2	3	3
G3-12	4	2	2	2
G3-13	1	1	0	0
G3-14	2	2	0	0
G3-15	2	1	0	0
G3-16	3	2	3	3
G3-17	3	2	3	3
G3-18	2	1	0	0
G3-19	2	1	1	1
Total	50	32	25	20
Average	2.63	1.68	1.32	1.05

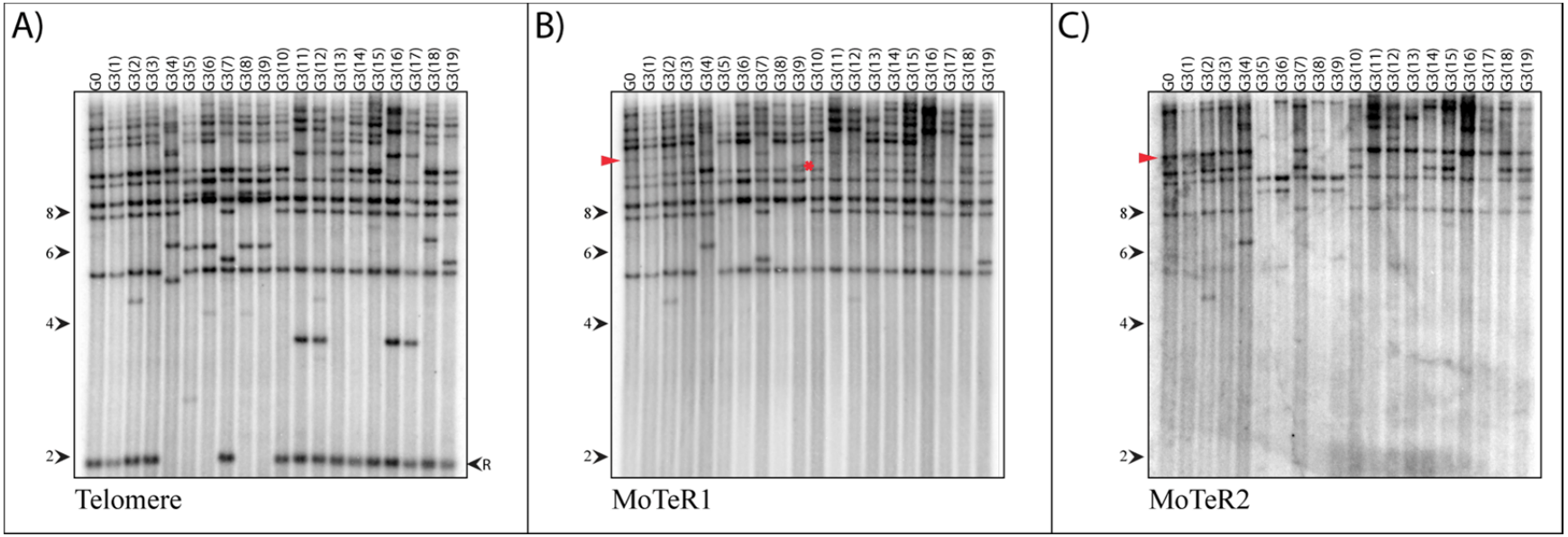

Figure 5-1. Southern analyses of LpKY97-1A mitotic progeny using telomere, MoTeR1, and MoTeR2 probes. Genomic DNA was digested with *Pst*I, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized with a telomere probe (A), a MoTeR1 probe (B), and a MoTeR2 probe (C). The blots were stripped between different probes. Black arrows indicate the size markers in kilobases. Red arrows to the left side of images indicate a MoTeR band that did not cohybridize with a telomere band. Lane 1 is the initial starting culture DNA fingerprint (G0). In lanes 2-20 are different third generation mitotic progeny of LPKY97-1A. The letter R represents the telomeric rDNA fragment in LpKY97-1A. Red asterisk indicates a MoTeR1 fragment that failed to cohybridize with telomeric fragment in the mitotic progeny G3(9).

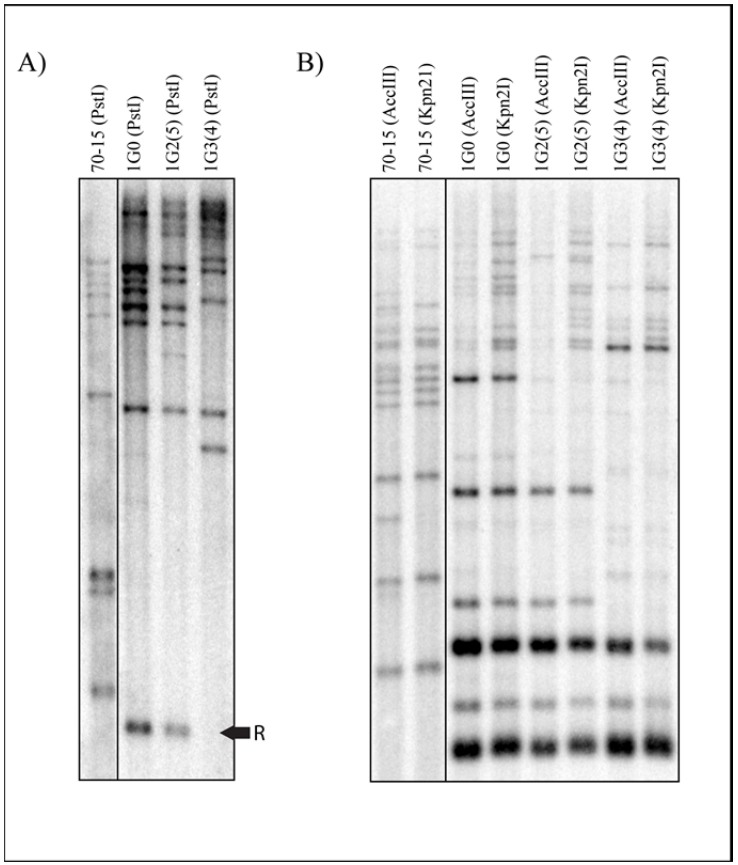

Figure 5-2. Methylation in telomere restriction fragment length polymorphisms. In A) the genomic DNA of four different isolates were digested with *Pst*I (sensitive to methylation). In lane 1 is the Oryza pathotype isolate 70-15, in lane 2-4 are genomic digests of different generations of the Lolium pathotype isolate LpKY97-1A. In B) lanes 1,3,5,7 the genomic DNA of the isolates were digested with *Acc*III (not affected by cytosine methylation), and in lanes 2,4,6,8 the genomic DNA of the isolates were digested with *Kpn*2I (cytosine methylation sensitive). Genomic digests were electrophoresed in a 0.7% agarose gel, electroblotted to a nylon membrane, and hybridized with a telomere probe. The label R is the telomeric rDNA fragment in LpKY97-1A.

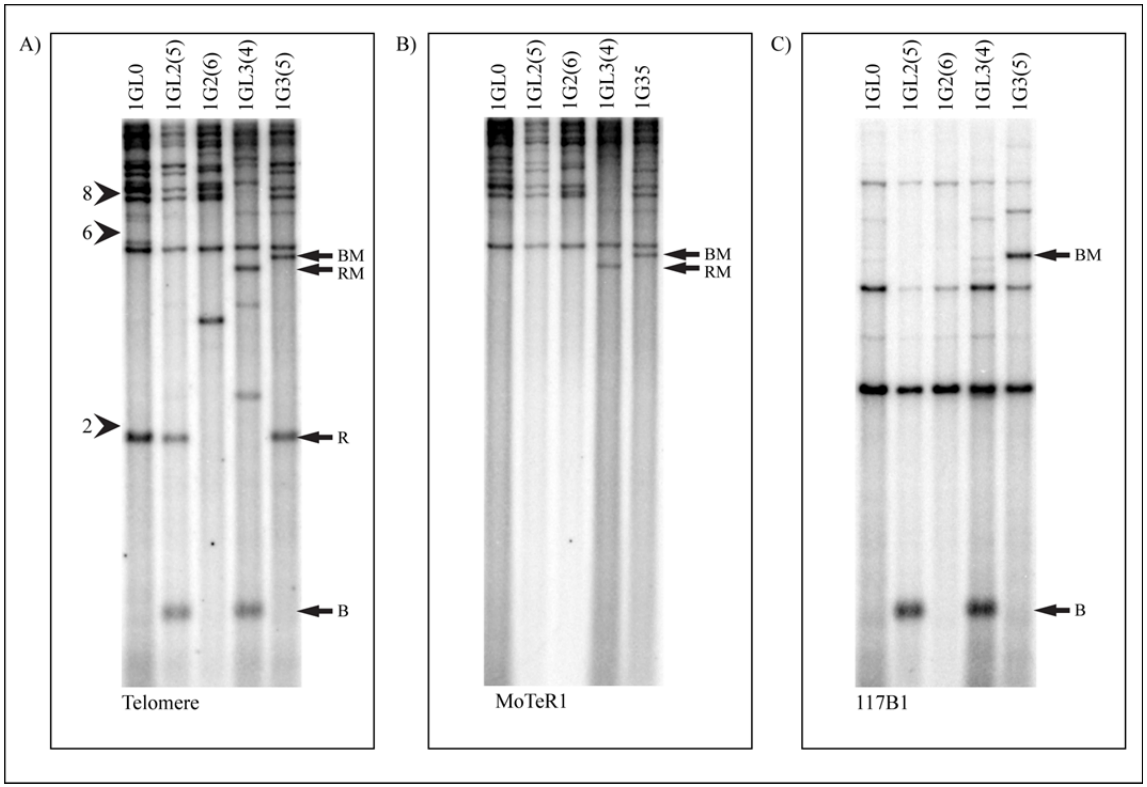

Figure 5-3. Characterization of telomere profile variation in mitotic progeny in LpKY97-1A. Genomic DNA of the isolates were digested with *Pst*I, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized sequentially with A) Telomere, B) MoTeR1, and C) 117B1. Blots were stripped between hybridizations. In lane 1 is the original culture of LpKY97-1A. In lanes 2-5 are mitotic progeny of LpKY97-1A with variant telomere profiles. Arrows indicate the size in kilobases. The label R is the telomeric rDNA fragment LpKY97-1A. The label RM is a variant version rDNA fragment that cohybridizes to MoTeR1. The label B is a *de novo* telomere fragment that corresponds to telomere 117B1. The label BM is another *de novo* telomere fragment that corresponds to the 117B1 telomere, but this fragment also cohybridizes with MoTeR1.

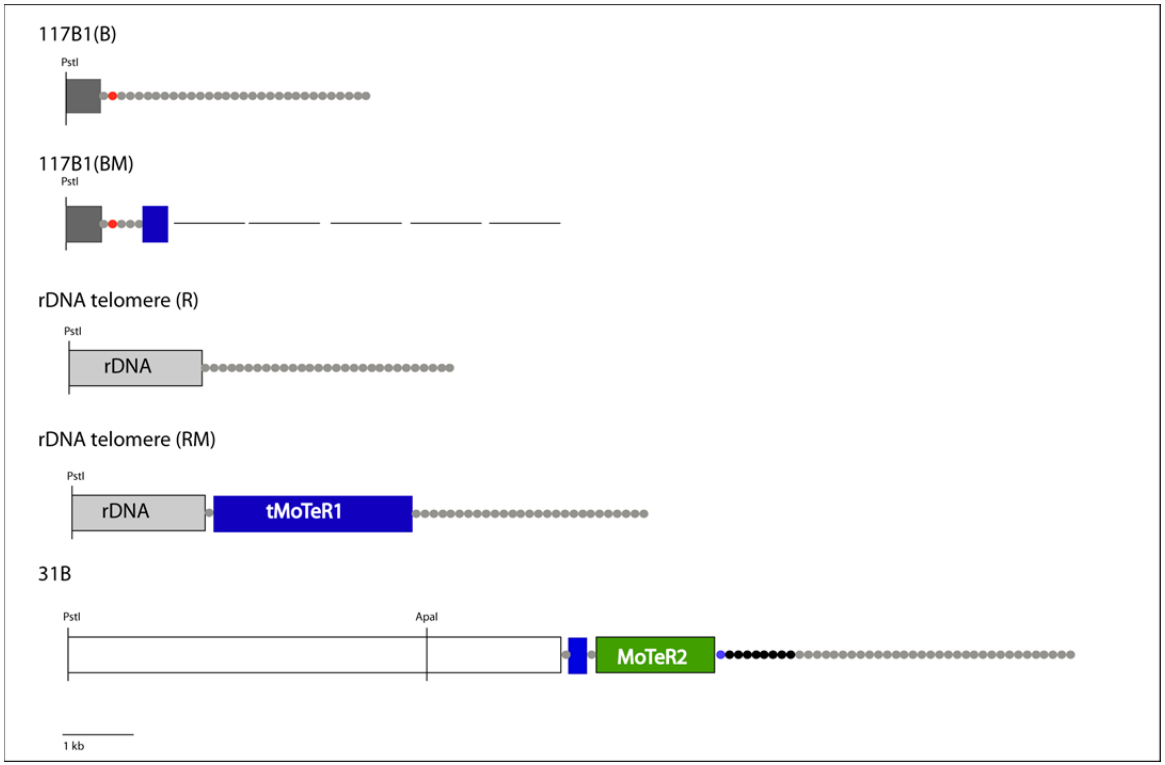

Figure 5-4. Characterized telomeres in LpKY97-1A. Graphical representation of telomeres captured in the end-enriched cloning process or through PCR amplification of specific junctions. Drawn to scale with the exception of telomere repeats and telomere like repeats that are expanded to show the structure of these repeats. Light grey circles represent the TTAGGG telomere repeat. Red circles represent the sequence TTAGG. Blue circles represent the sequence TTCGGG. Black circles represent the sequence TTTGGG. Blue rectangles represent MoTeR1. Green rectangles represent MoTeR2. Dark grey rectangles represent the sequence captured in clone 117B1. Light grey rectangles represent the sequence matching rDNA. White box represents unique sequence from the chromosome end 31B. Dashed line represents that the telomere end was not sequenced or captured in the end-enriched cloning process. The letters in parenthesis correspond to R = ribosomal, B = 117B1, M = MoTeR1.

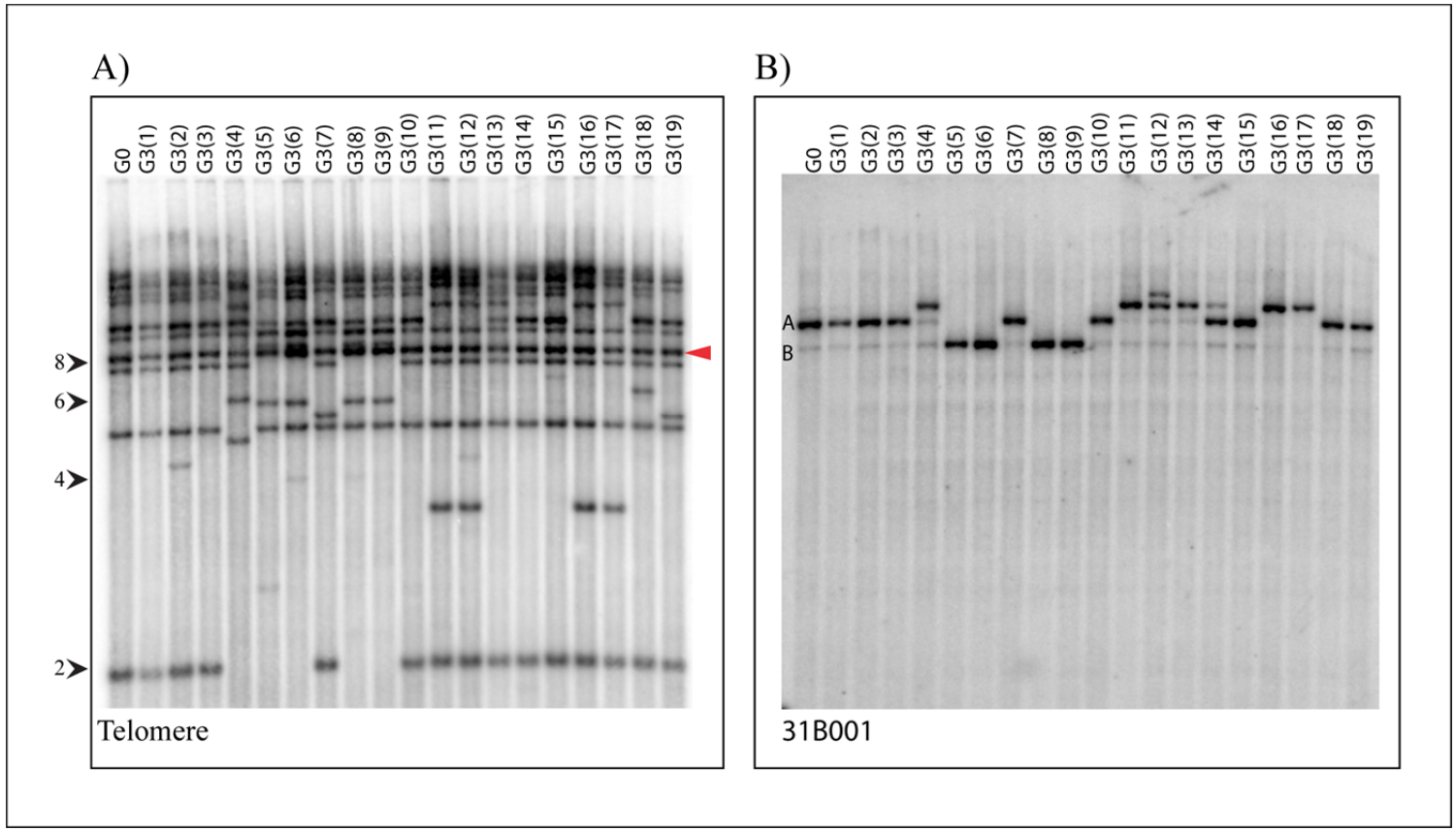

Figure 5-5. Southern analyses to detect changes at the 31B chromosome end in mitotic progeny of LpKY97-1A. Genomic DNA were digested with *Pst*I. Genomic digests were electrophoresed in a 0.7% agarose gel, electroblotted to a nylon membrane, and hybridized with a telomere probe (A) and the 31B001 probe (B). The membrane was then exposed to phosphorimaging screens and the resulting phosphorimages are shown. Membrane was stripped between probings. Lane 1 is the initial starting culture DNA fingerprint (G0). Lanes 2-20 are different third generation mitotic progeny of LPKY97-1A.

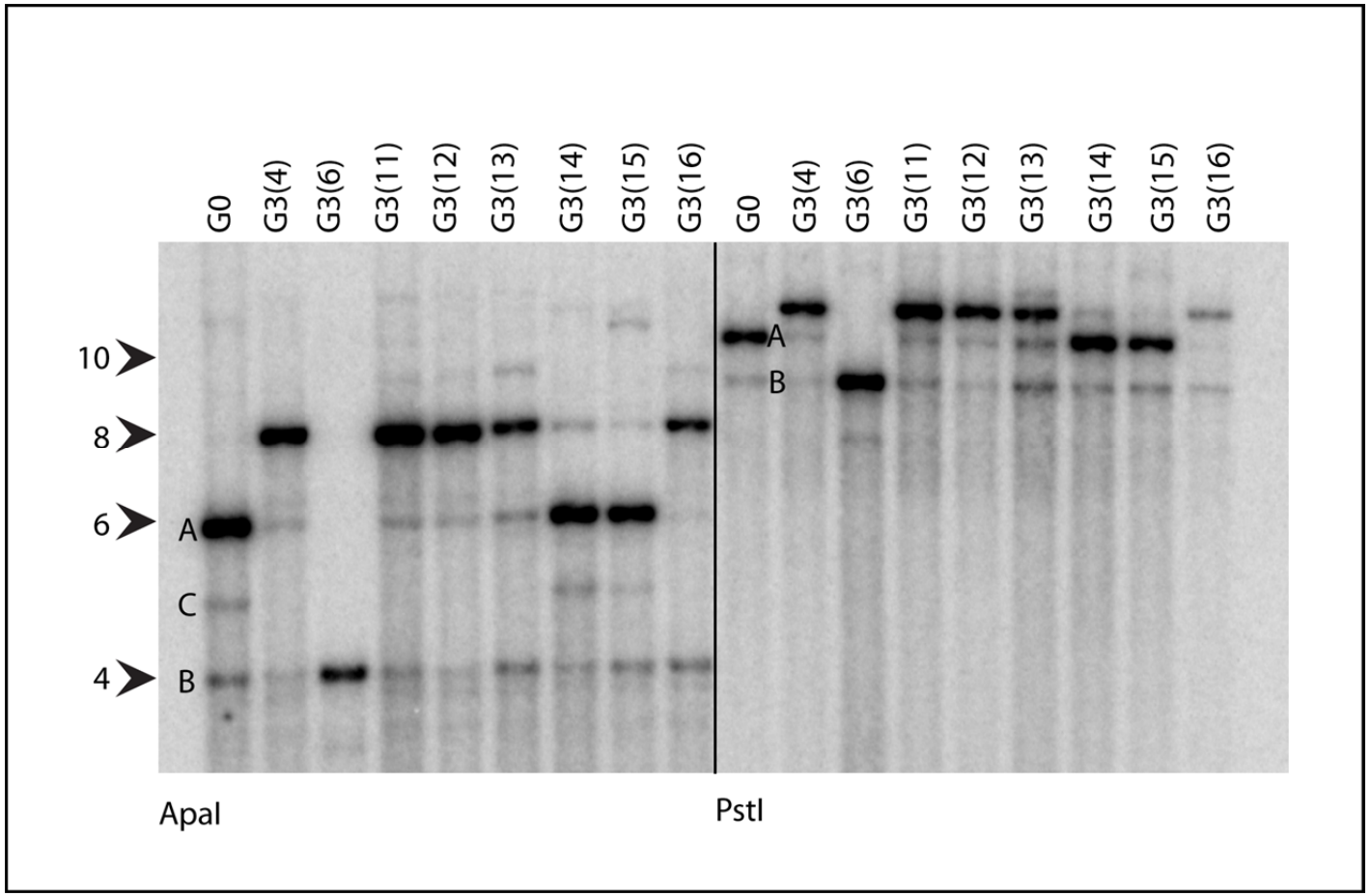

Figure 5-6. Resolution of 31B fragments in Southern analyses using either *Apa*I or *Pst*I digested genomic DNA. Single spores from generation 3 showing variation in 31B001 profile were digested with *Pst*I and *Apa*I and probed with 31B002. This was used to determine if probe 31B002 and *Apa*I digests would show the same profile, but yield smaller size bands. Arrows indicated size in kilobases. Genomic digests were electrophoresed in a 0.7% agarose gel, electroblotted to a nylon membrane, and hybridized with a telomere probe. The membrane was then exposed to a phosphorimaging screen and the resulting phosphorimage is shown.

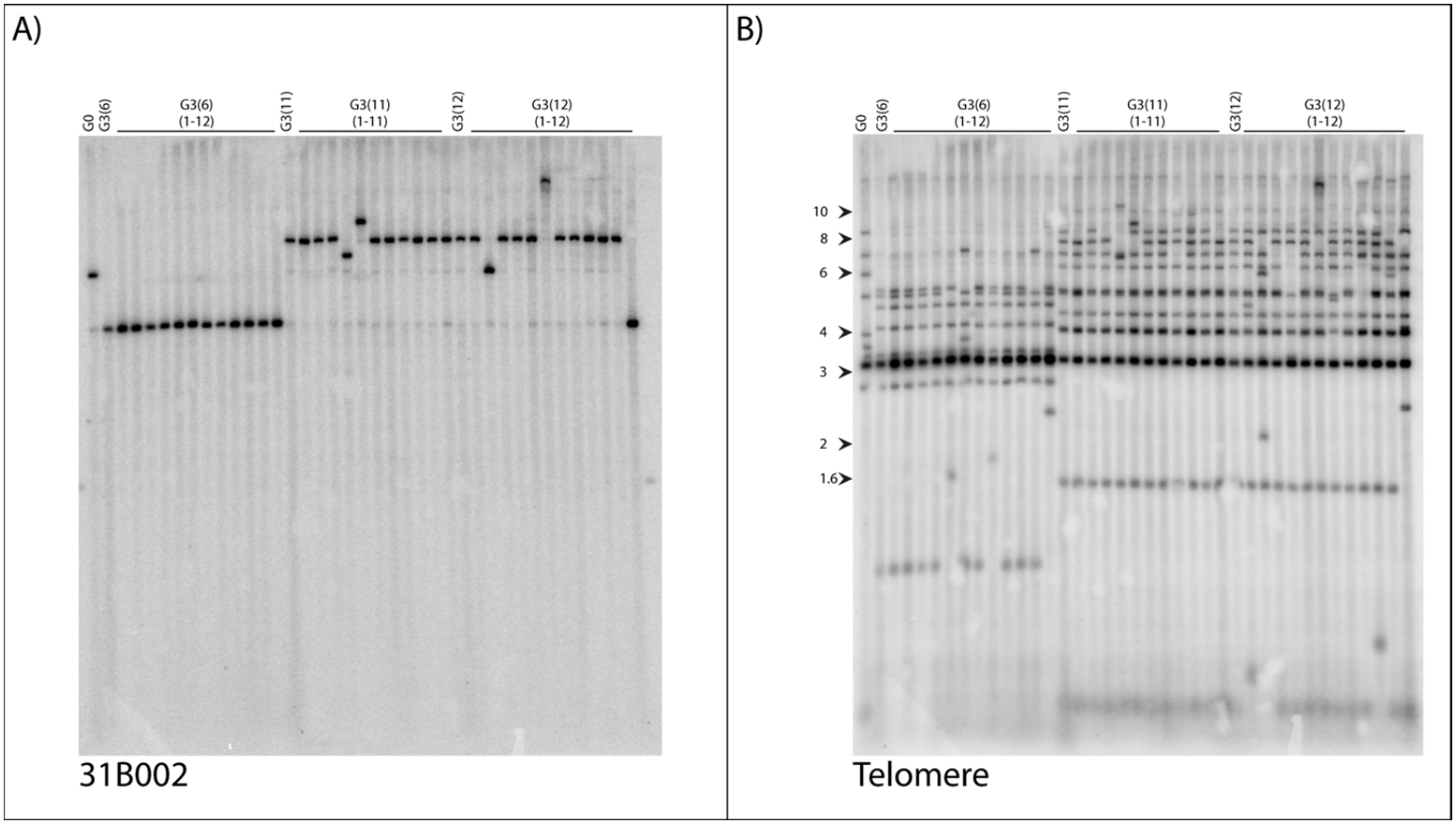

Figure 5-7. Southern analyses of progeny from three third generation variants of telomere 31B. LpKY97-1A progeny with variant 31B telomere profiles were single spored an additional round, DNA was digested with *Apa*I, electrophoresed in a 0.7% agarose gel, electroblotted to a nylon membrane, and hybridized with 31B002 probe (A) and telomere probe (B) to test for stability of the different telomere band sizes. The membrane was exposed to phosphorimaging screens and the resulting phosphorimages are shown. Membrane was stripped between probing. Arrows indicate size in kilobases. Lane 1 is the initial starting culture of LpKY97-1A. Lanes 2 (G3(6)), 15 (G3(11)), and 27 (G3(12)) are the third generation cultures that were further cultured to produce mitotic progeny shown under the solid lines.

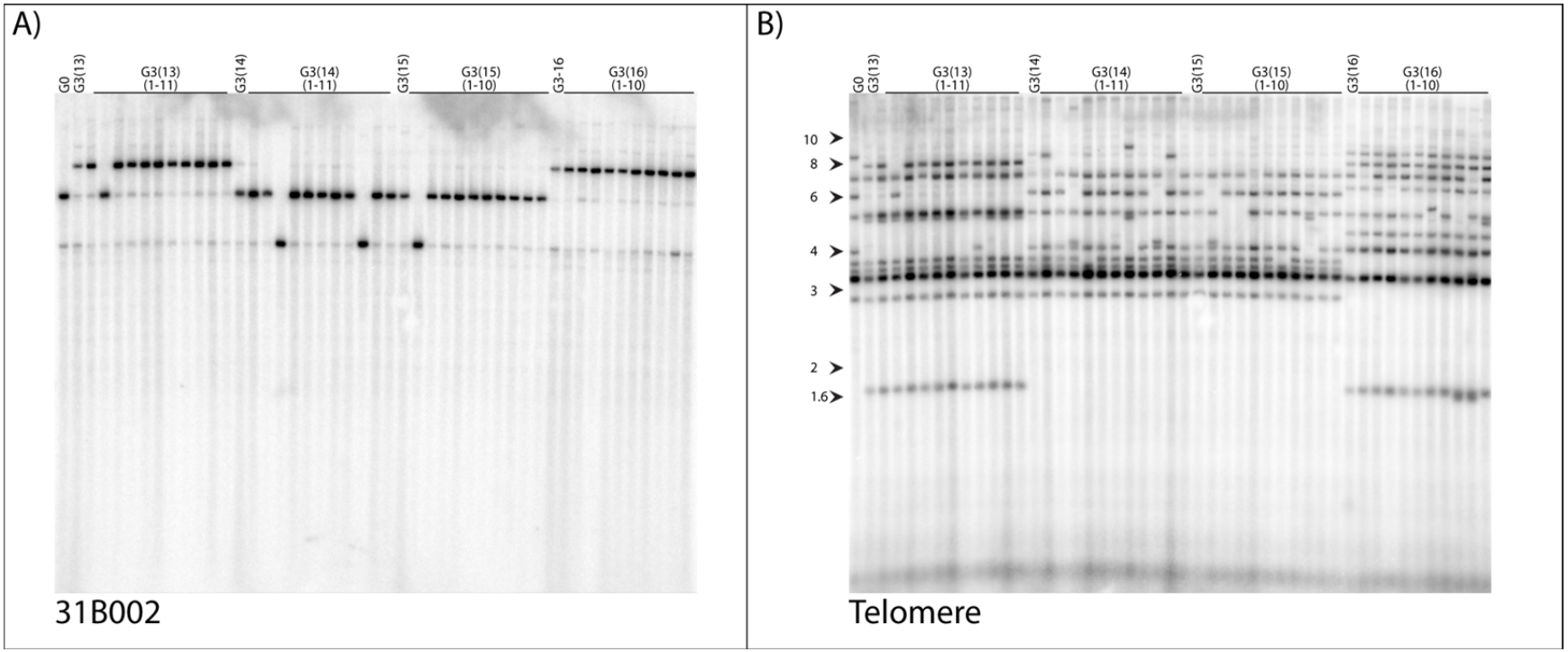

Figure 5-8. Sothern analyses of progeny from four third generation variants of telomere 31B. LpKY97-1A progeny with variant 31B telomere profiles were single spored an additional round, DNA was digested with *Apa*I, electrophoresed in a 0.7% agarose gel, electroblotted to a nylon membrane, and hybridized with 31B002 probe (A) and telomere probe (B) to test for stability of the different telomere band sizes. The membrane was exposed to phosphorimaging screens and the resulting phosphorimages are shown. Membrane was stripped between probings. Arrows indicate size in kilobases. Lane 1 is the initial starting culture of LpKY97-1A. Lanes 2 (G3(13)), 14 (G3(14)), 26 (G3(15)) and 37 (G3(16)) are the third generation cultures that were further cultured to produce mitotic progeny shown under the solid lines.

CHAPTER SIX

Distribution and Evolutionary History of *Magnaporthe oryzae* Telomeric Retrotransposons within Magnaporthe

6.1 Introduction

Magnaporthe oryzae includes several host-specialized subgroups, each of which are restricted in the range of host species in which they can cause disease. Representative subgroups are the *Eleusine* pathotype pathogenic on finger millets (*Eleusine coracana*); the *Lolium* pathotype pathogenic on perennial ryegrass (*Lolium perenne*), tall fescue (*Lolium arundinaceum*), and annual ryegrass (*Lolium multiflorium*); the Oryza pathotype pathogenic on rice (*Oryza sativa*); the Panicum pathotype pathogenic on torpedo grass (*Panicum repens*); the Setaria pathotype pathogenic on foxtails (*Setaria spp.*); and the Triticum pathotype pathogenic on wheat (*Triticum aestivum*) (KATO *et al.* 2000; TOSA *et al.* 2004). *Magnaporthe grisea* isolates, close relatives of *M. oryzae*, are virulent on crabgrass (*Digitaria spp*.).

Restriction length polymorphism (RFLP) analysis is one method that can be used to delineate isolates into their respective pathotypes. The unique banding pattern of different isolates in Southern hybridizations is referred to as the isolates "DNA fingerprint". DNA fingerprinting has been used to support the grouping of isolates into host-specific forms (BORROMEO *et al.* 1993; HAMER *et al.* 1989a; KATO *et al.* 2000). Repetitive elements have been used in DNA fingerprinting to resolve the association of *M. oryzae* isolates to the limited subset of cultivars in which they are virulent (CORREA-VICTORIA *et al.* 1994; LEVY *et al.* 1991). Two repetitive elements (MGR586 and MGR583), extensively used in DNA fingerprinting studies, contain sequence relating to different types of transposable elements (PoT3 and MGL

respectively) (FARMAN *et al.* 1996a; KACHROO *et al.* 1997). The distribution of transposable elements has frequently been used to understand population structure within and among *M. oryzae* pathotypes (KUMAR *et al.* 1999; PARK *et al.* 2008; ROUMEN *et al.* 1997; TANAKA and NAKAYASHIKI 2009; YAMAGASHIRA *et al.* 2008). The use of multiple transposable elements in repetitive DNA fingerprinting is attractive due to the differing evolutionary history for each transposable element.

MoTeRs are present in the genome of the Lolium pathotype isolates and in a single Eleusine pathotype isolate (PH42), but were absent in the laboratory Oryza pathotype isolate 70-15 (Figure 3-7). MoTeRs were amplified to a different degree within isolates of the Lolium pathotype. These results indicated that MoTeRs might be unevenly distributed in the Magnaporthe species. This was surprising as the other described non-LTR retrotransposon (MGL) in *M. oryzae* is present in most, if not all, host specialized pathotypes (ETO *et al.* 2001). Non-LTR retrotransposons tend to be inherited by vertical descent (MALIK *et al.* 1999), so one might expect MoTeRs to be present in most host specialized forms of Magnaporthe. The absence of MoTeRs in the widely studied laboratory isolate 70-15 necessitated the need to explore the overall distribution of MoTeRs in various pathotypes. Repetitive DNA fingerprinting was used to examine the distribution of MoTeRs in Magnaporthe isolates from 14 different pathotypes. Using information gathered from the distribution data, partial sequencing of MoTeR1s from different pathotypes, and evolutionary analysis of these sequences the likely evolutionary history of MoTeRs was addressed.

6.2 Results

6.2.1 Distribution of MoTeRs within Magnaporthe

To determine the distribution of MoTeRs in Magnaporthe, sequence-specific MoTeR1 and MoTeR2 probes were used as probes for Southern hybridization analyses of *Pst*I-digested DNA from 116 isolates representing 14 different pathotypes. Since some telomeric MoTeRs were highly truncated in Lolium pathotype isolates (Figure 3-4), two different MoTeR1 probes were utilized to distinguish between putatively intact and truncated MoTeR1s. The MoTeR1(5'RT) probe was used to identify presumed intact copies that retained the reverse transcriptase needed for transposition, while the MoTeR1(3') probe was used to detect truncated copies of MoTeR1. Since MoTeRs were mostly found within the telomeres of the Lolium pathotype isolates, I wanted to ascertain if MoTeRs were also telomeric in other pathotypes. To accomplish this, all of the Southern blots were subsequently hybridized with a telomere probe. Hybridizing fragments were counted as single bands regardless of intensity of the hybridization. The results of these experiments are summarized in Table 6-1 and Table 6-2.

Isolates from the Lolium pathotype all had a putatively intact MoTeR1 copy present within their genome (Table 6-1). There were more MoTeR1(RT) bands observed on average in the Lolium pathotype (~8 bands per isolate) than in other pathotypes. In addition to the Lolium pathotypes, other pathotypes had multiple apparently intact MoTeR1 in their genome. These included isolates from Triticum (Figure 6-1), Eragrostis (Figure 6-4 lanes 4 and 8) and two out of the four Eleusine (Figure 6-4 lanes 3 and 7) pathotype isolates. At least one putative intact copy of

MoTeR1 was present in all isolates surveyed from the Stenotaphrum, Paspalum, Leptochloea, and Zingiber pathotypes. In contrast, isolates from Setaria and Oryza showed uneven distribution of MoTeRs among isolates, with some showing a single intact copy of MoTeR1 while others were apparently completely devoid of the element. Finally, some pathotypes lacked full length MoTeR1s; these included Brachiaria, Panicum, Pennisetum, and Digitaria. Most of the putatively intact MoTeR1 bands in the various pathotypes were telomeric in nature (337/347 or 97%) (Table 6-2) based on cohybridization of MoTeR1(RT) probe and telomere probe to *Pst*I fragments in Southern analysis. Only isolates from the Lolium and Eleusine pathotypes had copies of intact MoTeR1s that were not telomeric in nature (6 and 4 fragments respectively). These fragments were likely once telomeric in nature.

The Southern analyses with the MoTeR1(3') probe, capable of detecting truncated MoTeR1 copies, provided more thorough analyses of MoTeR1 presence/absence within the different pathotypes. According to this analysis, only the Panicum and Pennisetum pathotypes completely lacked MoTeR1. At least one partial copy of the element was present in all isolates surveyed from Lolium, Eragrostis, Eleusine, Triticum, Leptochloea, Zingiber, Paspalum, and Setaria. Truncated MoTeR1 copies still exhibited a spotty distribution within the Oryza pathotype. Isolates from the Digitaria pathotype, members of the sister species *M. grisea*, all had MoTeR1 present within their genomes (Figure 6-5). Only 103 of the 185 (or 56%) truncated MoTeR1 bands were telomeric, which is lower than the percentage of apparently intact MoTeR1 bands that were associated with telomeres

(97%) (Table 6-2). Truncated MoTeR1 thus have a substantial decrease in telomeric association. It is likely that these truncated MoTeR1s have become internalized in the genome.

MoTeR2 was detected in 42 of the 116 isolates analyzed (Table 6-1). It was restricted to isolates that also had MoTeR1 in their genome. Most Lolium pathotype isolates tested (25/27) had MoTeR2s in their genome. All of the Eleusine pathotype isolates had MoTeR2s. MoTeR2 was detected in only half of the Triticum pathotype (9/18) and Eragrostis pathotype isolates (1/2). Only two isolates from the Setaria pathotype had MoTeR2. In an Oryza pathotype laboratory strain (2539), a single MoTeR2 band was observed in the Southern hybridizations (Figure 6-3, lane 5). DNA fragments of isolates surveyed from the Brachiaria, Digitaria, Leptochloea, Panicum, Paspalum, Pennisetum, Stenotaphrum, and Zingiber pathotypes did not hybridize to MoTeR2 probe in the Southern analysis (Figure 6-4 and Figure 6-5), likely indicating that they lacked MoTeR2, or that the remaining fragments were severely truncated. Most of the MoTeR2 fragments in the Southern analysis cohybridized with telomere fragments (200/202) (Table 6-2), indicating that they were probably telomeric. The only exceptions to this were in two Setaria pathotype isolates (YF1 and YF2).

6.2.2 MoTeR evolution within *Magnaporthe oryzae*

MoTeR1 was present in most pathotypes of Magnaporthe. Multiple putatively intact copies of MoTeRs were also present in Eleusine and Eragrostis pathotypes rooted deeply in the proposed evolutionary history of *M. oryzae* (COUCH *et al.* 2005). This suggested that MoTeR1 was inherited by vertical descent and, therefore, was lost in

some isolates. In this case, it would be expected that the evolutionary history of MoTeR1 would resemble those of other nuclear genes within the organism. To test this, three genomic loci (MPG1, CH7-Bac7, and the rDNA internal transcribed spacer [ITS] region) were amplified by PCR from genomic DNA in 26 different MoTeR1 containing isolates and the PCR amplicons were cloned and sequenced. Likewise, MoTeR1 sequences in the various isolates were amplified by PCR from genomic DNA and resulting PCR amplicons were then cloned and sequenced. The sequences were aligned and the evolutionary history was inferred using the neighbor-joining (NJ) method (SAITOU and NEI 1987). The bootstrap tree was derived from 1000 replicates and is taken to represent the evolutionary history of the 26 taxa analyzed (FELSENSTEIN 1985). Missing data or alignment gaps were eliminated only in pairwise sequence comparisons.

The data from individual gene loci provided very little phylogenetic resolution among pathotypes due to some isolates sharing identical sequence. Data from all gene loci were concatenated to improve resolution. A neighbor-joining tree based on 1323 positions in the concatenated sequence of the genomic loci is shown in Figure 6-6-A. The neighbor-joining phylogenetic tree of MoTeR1 based on 1373 positions is shown Figure 6-6-B. The topology of the gene loci and MoTeR1 trees were compared to determine if there was a similar evolutionary history.

According to the NJ analyses of the gene loci, isolates within pathotypes were grouped together with the exception of the Lolium pathotype (Figure 6-6-A). The Lolium(1) pathotype isolates (LpKY97-1A, FH, NJ1, and NJ2) were all collected in the United States, while the Lolium(2) pathotype isolate (WK31) was a weakly

aggressive isolate that was collected in Japan. Lolium(1) pathotype isolates were most similar to Eleusine pathotype isolates(CD156, PH42, JP29, and G22), while the Lolium(2) pathotype isolate was more similar to Triticum (BR2 and BR4), Setaria (MCCL2, ARC1, STHLAND, YFDC, and GF1), and Oryza pathotype isolates (2539, 91T14, and IT11). The gene loci delineated the Eleusine pathotypes into two groups based on the species of hosts they were collected from: *Eleusine coracana* (JP29 and G22) and *Eleusine indica* (CD156 and PH42). Stenotaphrum isolates (PG1108, PG1054, SSFL, and STAGMS) were most similar to Eragrostis pathotype isolates (G17 and AR4) based on gene loci.

The evolutionary history of MoTeR1 appeared to be somewhat different than the evolutionary history of gene loci (Figure 6-6-B), and there were a higher number of base substitutions in the MoTeR1 sequences than were calculated in the nuclear genes (Figure 6-6). In the previous NJ analysis of nuclear genes the Triticum pathotype isolates were more closely related to Lolium(2), the Setaria, and Oryza isolates. However, the MoTeR1s of Lolium(1) pathotypes were highly similar to the MoTeR1 sequences of Triticum pathotype isolates. The Setaria isolates were all grouped together based on nuclear genes, but the MoTeR1 of ARC1 was most similar to the MoTeR1s of Lolium(1) and Triticum pathotype isolates. The Eleusine pathotype was polyphyletic in regards to MoTeR1, but monophyletic based on nuclear genes. The MoTeR1s of Eleusine indica isolates (CD156 and PH42) were similar to the MoTeR1 sequence in the Lolium(2) (WK31) isolate and two Eragrostis pathotype isolates (G17 and AR4). The MoTeR1s of *Eleusine coracana* pathotype isolates (G22 and JP29) were similar to most of the MoTeR1s in Setaria pathotype

isolates (YFDC, MCCL2, STHLAND). The Oryza pathotype isolates' MoTeR1s grouped together with the exception of 2539, a laboratory strain, where the MoTeR1 sequence grouped with the MoTeR1 of Eragrostis isolate G17. The Stenotaphrum pathotype isolates were grouped together in regards to MoTeR1, and were closely related to the MoTeR1s of the Oryzae pathotype isolates. This was different from what had been observed previously in the NJ analysis of nuclear genes where Stenotaphrum isolates were more closely related to Eragrostis isolates.

One factor that could lead to grouping of MoTeR1s in phylogenetic trees is that when analyzing the sequence of MoTeR1 in an individual pathotype the orthologous copies were compared, while paralogous MoTeR1 copies might have been compared in the analyses between different pathotypes. For example, isolates IT11 and 91T14 were closely related phylogenetically (Figure 6-6-B) and produced a MoTeR1 band with a similar size in Southern analysis (Figure 6-3-B). The SIT4 isolate had a MoTeR1 band that was smaller in size. This could suggest it was a truncated form of MoTeR1 on the same chromosome end as 91T14, or it could represent a paralogous MoTeR1 copy on a different chromosome end. To determine whether the MoTeR1 in 91T14, IT11, and SIT4 were on the same chromosome end, Josh Moore, an undergraduate in Dr. Mark Farman's lab, developed a probe that was specific to the MoTeR1 containing chromosome in IT11. I hybridized this probe with the Southern blot shown in Figure 6-3. Only one band in 91T14 and IT11 cohybridized to the telomere and the MoTeR1(5'RT) band that was identified earlier (Figure 6-3-E). This suggested that MoTeR1 was indeed at the homologous chromosomal end in these two isolates. The IT11 probe did not hybridize with the

band that hybridized to MoTeR1(5'RT) probe in SIT4 DNA (Figure 6-3-E), which suggested that the intact MoTeR in SIT4 was on a different chromosome end. Since the SIT4 MoTeR1 is on a different chromosome end, this copy is likely paralogous to the MoTeR1 copy in the IT11 and 91T14 genome. The MoTeR1 sequences derived by PCR of genomic DNA from IT11, 91T14, and SIT4 were highly similar and grouped together in phylogenetic analysis even though the SIT4 MoTeR1 copy was a paralog. This suggests that the effect of using sequence derived from orthologous versus paralogous MoTeR1 copies was not a likely factor in the grouping of the MoTeR1s in the NJ tree.

6.2.3 Possible mechanisms affecting the evolutionary history of MoTeRs

MoTeRs had more mutations between the various pathotypes than were observed in the nuclear genes. The repeat-induced point mutation mechanism (RIP) can promote accelerated evolution and is known to occur in Magnaporthe (IKEDA *et al.* 2002). RIP causes nucleotide transitions whereby G:C are changed to A:T in duplicated DNA elements during the sexual reproductive phase (CAMBARERI *et al.* 1989). Thus it would be expected that MoTeR1 sequences would show a strong bias towards transitions mutations if RIP were active. A maximum composite likelihood estimate of the pattern of nucleotide substitution was calculated between the 26 isolates examined in the neighbor-joining analyses to determine if a bias in transition versus transversions existed. The rates of a transitional mutation were higher in each base observed (Table 6-3). The nucleotide frequencies were 0.359 (A), 0.242 (T), 0.225 (C), and 0.173 (G). The transition/transversion rate ratios were k_1=3.623 (purines) and k_2=7.335 (pyrimidines). The overall

transition/transversion bias was R=2.042, where R=[A*G*k_1 + T*C*k_2]/[(A+G)*(T+C)]. An overall transitions/transversion bias of 2.042 is not typically indicative of RIP'd sequences. To examine the possibility of RIP more closely, MoTeR1 sequence from LpKY97-1A (multiple copies of MoTeR1) and IT11 (a single copy of MoTeR1) were aligned using MUSCLE, and the sequences were compared in MEGA version 4 (See Appendix C for data). A total of 1374 bp were in the alignment with two indels. After removing the indels from analyses, the sequence was highly conserved between LpKY97-1A and IT11 (1285 out of 1364 bases conserved or 94%). There were a total of 79 nucleotide differences in MoTeR1 sequence between LpKY97-1A and IT11. Transitions made up 52 out of the 79 nucleotide differences between the two sequences, while there were 27 transversions. The ratio of transitions to transversions was 1.93:1. The background transition to transversion ratio in Magnaporthe is 2:1 (COUCH and KOHN 2002). Since RIP didn't appear to be the predominant mechanism for the sequence difference in MoTeR1 between host pathotypes, it seems more likely that random mutation may have led to the sequence divergence between different host pathotypes.

6.3 Discussion

6.3.1 Distribution of MoTeRs in Magnaporthe

The major objectives of the experiments in this chapter were to determine the distribution of the MoTeR retrotransposons within *Magnaporthe,* and based on these data, to describe the possible evolutionary history of MoTeRs. In previous chapters, it was noted that MoTeR1s were present in isolates of the Lolium pathotype, but absent in isolates of the Oryza pathotype. It was therefore predicted

that MoTeRs would be unevenly distributed among the many different host specific isolates within *M. oryzae*. This initial observation was confirmed through further experimentation. The Lolium pathotype had the most MoTeR1 copies followed by the Triticum pathotype isolates. Only a few Oryza pathotype isolates had MoTeR1 elements. The uneven distribution of MoTeR1 was surprising because the other major non-LTR retrotransposon in *M. oryzae* was more widely distributed (ETO *et al.* 2001). Several processes could be involved in the differences in copy number of MoTeR1 such as stochastic loss, an increased rate of retrotransposition, inactivation of repetitive sequences either by passive or active processes, horizontal transmission, or self-regulation of transposition (ARKHIPOVA 2005; JOHNSON 2007; LE ROUZIC and CAPY 2005).

MoTeR2 distribution was more limited than MoTeR1 and primarily confined to isolates that had MoTeR1. MoTeR1 was typically found at a higher copy number in the isolates that also had MoTeR2. Based on the prediction that MoTeR2 requires MoTeR1 for transposition, only the horizontal transmission of a MoTeR1 or both MoTeR1 and MoTeR2 would allow for amplification in a new genome. It is not yet known whether a MoTeR1 element is capable of horizontal transmission. Also, it is not known whether MoTeRs would have an ability to amplify in the genomes of isolates that did not previously have MoTeRs. Previous work with the LTR retrotransposon MAGGY has shown that it can amplify in a genome not previously containing the element (NAKAYASHIKI *et al.* 2001a). MoTeRs have been shown to amplify under conditions where normal telomerase function has been abolished (Farman unpublished). This could suggest that MoTeRs are amplified in isolates

with a lower expression of telomerase or in isolates missing important regulators in telomere maintenance, which would be an interesting avenue of research to pursue.

This study represents one of the few studies where multiple probes were used from a single retrotransposon. In the case of the MAGGY and Grasshopper LTR retrotransposons, only probes that represented the internal portion of the element were used (DOBINSON *et al.* 1993; FARMAN *et al.* 1996b). There is a possibility that, like MoTeRs, the MAGGY and Grasshopper LTR retrotransposons were lost in some isolates while being retained in others. To determine whether degraded LTR elements exist within other *M. oryzae* genomes, further experimentation using probes from the LTR region of MAGGY and Grasshopper could be conducted.

6.3.2 Possible routes to the present evolutionary pattern of MoTeR1

There are two possible explanations to account for the uneven distribution of MoTeR1 in *M. oryzae*. The first explanation supposes that the MoTeRs are ancient retrotransposons that were present in the common ancestor but were absent in later generations due to either stochastic loss or by an active mechanism such as repeat induced point mutation (RIP). While a majority of the sequence differences between LpKY97-1A and IT11 were transition mutations, it is not as striking of a pattern that has been observed in sequences where RIP is active and nearly all the nucleotide substitutions are transitions (BRAUMANN *et al.* 2008). In fact, the transition/transversion ratio for MoTeR1 (1.93:1) is lower than previous studies that calculated the transition/transversion ratio among the gene loci in *M. oryzae* isolates (COUCH and KOHN 2002). The t/v ratio of other transposable elements in *M. oryzae* are typically higher than gene loci, such as MAGGY (6.8:1), MGL (4.2:1),

and Pot2 (2.3:1) (THON *et al.* 2004). The low t/v ratio in MoTeR1 between LpKY97-1A and IT11 suggests that RIP is not actively involved in the DNA sequence differences between different host specific isolates. However, this evidence is based on sequence divergence of a single MoTeR1 copy in a single Oryza pathotype isolate compared to a single MoTeR1 amplified from LpKY97-1A. If the Oryza pathotype isolate never had more than one MoTeR1 in its genome, then it wouldn't be expected for RIP to be active on this sequence because RIP acts on repetitive elements within the genome (CAMBARERI *et al.* 1989). A wider study examining sequence from multiple MoTeR1 copies in LpKY97-1A, or another isolate with multiple MoTeR1 copies, might show evidence of RIP'd copies, and thus RIP as a mechanism for loss of MoTeRs cannot be completely ruled out.

An alternative explanation for the evolution of MoTeRs is that they could have arisen by horizontal transmission from one host isolate to another in the recent past. Horizontal transmission is often invoked to describe inconsistencies in the presence of retrotransposons in the evolutionary history of an organism. More traditionally, horizontal transmission is implied when transposable elements are discontinuously distributed in distantly related taxa in a way that cannot be explained by vertical inheritance (HARTL *et al.* 1997; KIDWELL 1992). The limited distribution of MAGGY and Grasshopper have been used to suggest that they likely arose through horizontal transmission (DOBINSON *et al.* 1993; ETO *et al.* 2001; FARMAN *et al.* 1996b). Horizontal transmission of non-LTR retrotransposons is thought to be a rare occurrence (MALIK *et al.* 1999), and there are only a few possible cases of horizontal transmission of non-LTR retrotransposons (ZUPUNSKI *et al.* 2001). The

reason for this rarity is believed to be due to the mechanism of transposition itself. In LTR retrotransposons, a DNA intermediate is produced which then inserts into the genome during transposition (LUAN *et al.* 1993). In non-LTR transposition, a RNA intermediate is reverse transcribed and the cDNA integrates directly into the target site (LUAN *et al.* 1993). The horizontal transmission of the RNA intermediate of MoTeR1 followed by successful integration into the genome is not likely. LTRs are much more likely to be horizontally transferred due to a DNA intermediate which is packaged in virus-like particles (CHAPMAN *et al.* 1992; FENG *et al.* 2000; WILHELM *et al.* 1994). These factors would increase the probability of LTRs surviving horizontal transmission.

The Southern analyses with the MoTeR1(3') probe, which allows for the detection of truncated MoTeR1s, indicated that many of the isolates had MoTeR1 in their genomes. The truncated MoTeRs were also seen in *M. grisea*, which is a separate species from *M. oryzae*. These data favor the hypothesis that the MoTeRs are ancient retrotransposons that were present in a common ancestor to both *M. oryzae* and *M. grisea*. Since most of the copies of MoTeR1 in distantly related taxa could be a result of vertical descent, it is interesting that only some lineages retained the MoTeR1 in their genome.

The conclusion that MoTER1 had a predominately vertical mode of descent was somewhat confounded by sequence similarity of MoTeR1 in ARC1 (a Setaria pathotype isolate) to the MoTeR1s sequence of other Lolium pathotype isolates. ARC1's MoTeR1(5'RT) fragment observed in the Southern blot was stronger in intensity, as well as being smaller in size, than the fragments observed in the other

Setaria pathotype isolates (data not shown). Additionally, in a neighbor-joining tree of nuclear genes the ARC1 sequence was similar to the other Setaria pathotype isolates (Figure 6-6-A), while ARC1's MoTeR1 sequence was more similar to the MoTeR1s in the Lolium(1) pathotype. Horizontal transfer is commonly implied when the sequence of a retrotransposon in an individual is more similar to a distantly related group than the sequence similarity observed at other loci (SILVA *et al.* 2004). However, there is another plausible explanation to describe the sequence similarity of the MoTeR1 in ARC1 to MoTeR1 copies in the Lolium(1) pathotype that would not violate vertical descent. The MoTeR1 fragment in ARC1 could represent a copy of the element, which was lost in the other Setaria pathotypes isolates. This MoTeR1 was telomeric, which would make it more likely to be lost through degradation of the chromosome ends. Setaria isolates are known to undergo frequent rearrangements at the chromosome ends (Farman, unpublished data), making the loss of a telomeric MoTeR copy highly possible. The process described above would fit with the model that MoTeRs were once present in all genomes. Nevertheless, horizontal transmission cannot be ruled out in the explanation of the similarity of MoTeR1 sequence in ARC1 to the MoTeR1 sequence of the Lolium pathotype isolates. Detailed comparative genomic analyses may be required to determine the possible origin of the ARC1 MoTeR1 element.

The evolutionary history of MoTeR1 elements did not match the evolutionary history of nuclear genes in *M. oryzae* based on NJ analyses (Figure 6-6). Instead, the sequence of MoTeR1s in isolates with higher copy number of putatively intact MoTeR1s from the Southern analyses tended to be more similar with one another. It

also held true that most pathotypes grouped together based on NJ analysis of MoTeR1 sequence with the exception of ARC1, discussed above, and the Eleusine pathotype isolates. The separation of Eleusine pathotypes into two distinct clades has also been observed in evolutionary analyses using other transposable elements (TANAKA and NAKAYASHIKI 2009). The MoTeR1s of the Eleusine pathotype isolates PH42 and CD156 were similar to the MoTeR1s of the Lolium pathotype, and they also had more copies of MoTeR1 in their genomes than G22 and JP29, which were more closely related to Setaria isolates based on MoTeR1 sequence similarity. The Oryza field isolates grouped closely to one another, although the isolates were separated by vast geographical distance and, in the case of SIT4, the MoTeR1 was likely a paralogous copy of MoTeR1 in regards to the other Oryza pathotype isolates. The Oryza pathotype laboratory strain 2539 was closely related to the Eragrostis pathotype isolates, which was anticipated as this isolate was a sexual progeny produced in crosses between Eragrostis and Oryza pathotype isolates. The MoTeR1 sequences in the Stenotaphrum pathotype isolates were similar to MoTeR1 sequences of Oryza pathotype field isolates, and formed a clade in the neighbor-joining analysis (Figure 6-6 B). However, based on NJ analysis of nuclear genes the two pathotypes were not as closely related.

One possible explanation for the sequence similarity of MoTeR1s in isolates with putatively intact copies could be that isolates with active MoTeR1s constantly replenish the MoTeR1 at the chromosome ends with a functional copy. In isolates with a single MoTeR1 that copy could easily become non-functional and then the sequence would undergo further mutagenesis or degradation. Copy numbers of

non-LTRs are generally high in eukaryotic organisms, but usually there are only a few master copies that have the ability to give rise to new copies (BROUHA *et al.* 2003; ZAGROBELNY *et al.* 2004). These "functional" copies would need to persist in the genome in order to continually produce more MoTeRs as telomere truncations may arise. The likely evolutionary history presented by the neighbor-joining analyses could represent the approximate time since the transposable element was functional. Lolium and Triticum pathotype isolates both had a high copy number of the MoTeR1 elements in their genomes. In this situation, it would also be possible to maintain sequence similarity of MoTeR copies by gene conversion (KIJIMA and INNAN 2010).

Table 6-1. Hybridization signals of MoTeR probes in various host pathotypes of *Magnaporthe*

Host pathotype of Magnaporthe	Number of isolates tested	Number of isolates with hybridization signals to MoTER1(5'RT)	Number of isolates with hybridization signals to MoTER1(3')	Number of isolates with hybridization signals to MoTER2
Brachiaria	2	0	1	0
Digitaria	16	0	16	0
Eleusine	4	4	4	4
Eragrostis	2	2	2	1
Leptochloea	1	1	1	0
Lolium	27	27	27	25
Oryza	19	4	12	1
Panicum	1	0	0	0
Paspalum	1	1	1	0
Pennisetum	2	0	0	0
Setaria	18	6	18	2
Stenotaphrum	4	4	4	0
Triticum	18	18	18	9
Zingiber	1	1	1	0
Total	116	68	105	42

Table 6-2. Telomeric association of MoTeR hybridizing fragments within host pathotypes of Magnaporthe

Host	Telomeric MoTER1(5'RT) bands[A]/total bands (%)	Telomeric MoTER1(3') bands[B]/total bands(%)	Truncated MoTeR1 telomeric bands[C]/total bands (%)	MoTER2 telomeric bands/total bands (%)
Brachiaria	0/0	1/1(100)	1/1(100)	0/0
Digitaria	0/0	24/36(67)	24/36(67)	0/0
Eleusine	17/21(81)	28/37(76)	11/16(69)	15/15(100)
Eragrostis	8/8(100)	11/12(92)	3/4(75)	5/5(100)
Leptochloea	1/1(100)	1/1(100)	0/0	0/0
Lolium	211/217(97)	222/244(91)	15/31(49)	163/163(100)
Oryza	5/5(100)	9/16(56)	4/11(36)	1/1(100)
Panicum	0/0	0/0	0/0	0/0
Paspalum	1/1(100)	2/2(100)	1/1(100)	0/0
Pennisetum	0/0	0/0	0/0	0/0
Setaria	6/6(100)	27/66(40)	21/60(35)	0/2(0)
Stenotaphrum	4/4(100)	12/14(86)	8/10(80)	0/0
Triticum	83/83(100)	86/86(100)	14/14(100)	16/16(100)
Zingiber	1/1(100)	2/2(100)	1/1(100)	0/0
Total Bands	337/347(97)	425/517(82)	103/185(56)	200/202(99)

[A] MoTeR(5'RT) hybridizing fragments were considered as intact MoTeR1s in these isolates.
[B] All fragments which hybridized to the MoTeR1(3') probes were counted including ones that cohybridized with MoTeR1(5'RT)
[C] MoTeR1(3') fragments which did not cohybridize with MoTeR1(5'RT) were counted. These were considered truncated MoTeRs.

Table 6-3. Maximum composite likelihood estimate of the pattern of nucleotide substitution

	A	T	C	G
A	-	*3.3*	*3.06*	**8.51**
T	*4.87*	-	**22.44**	*2.35*
C	*4.87*	**24.23**	-	2.35
G	**17.65**	*3.3*	*3.06*	-

The numerical entry represents the probability of substitution from one base (row) to another base (column) instantaneously. Comparison should only be made between entries within a row. Rates of transitional substitutions are shown in bold, and transversional substitutions are shown in italics. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons. There were a total of 1378 positions in the dataset.

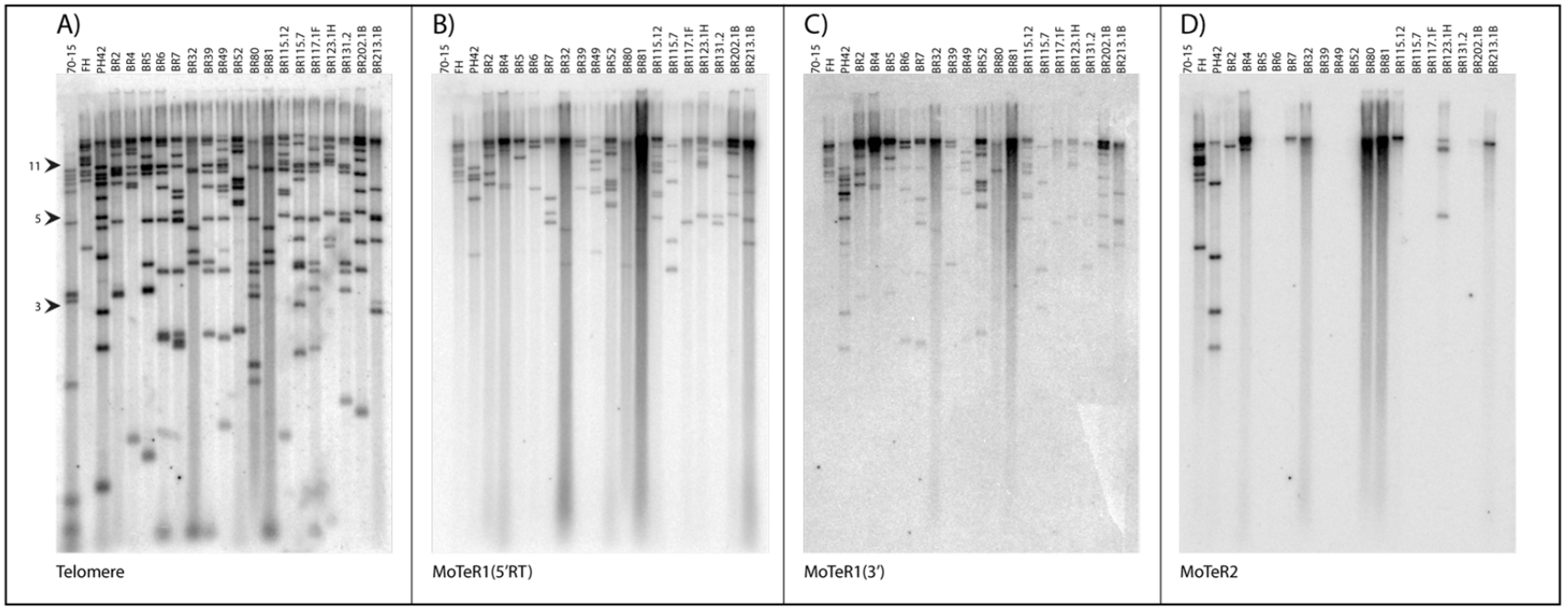

Figure 6-1. Southern analyses of genomic DNA of *M. oryzae* isolates from Triticum pathotype isolates. Genomic DNA was digested with *Pst*I, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized sequentially with A) Telomere, B) MoTeR1(5'RT), C) MoTeR1(3') and D) MoTeR2. The membrane was exposed to phosphorimaging screens and the resulting phosphorimages are shown. Membrane was stripped between probings. Lanes 1-3 are isolates used as controls 70-15 is the laboratory strain negative for the MoTeRs, FH was isolated from Lolium and is a positive control for MoTeRs, PH42 was isolated from Eleusine and is an additional isolate that is positive for MoTeRs. In lanes 4-21 are *M. oryzae* isolates that were collected from Triticum in Brazil. Black arrows to the left of the blot indicate the size markers in kilobases.

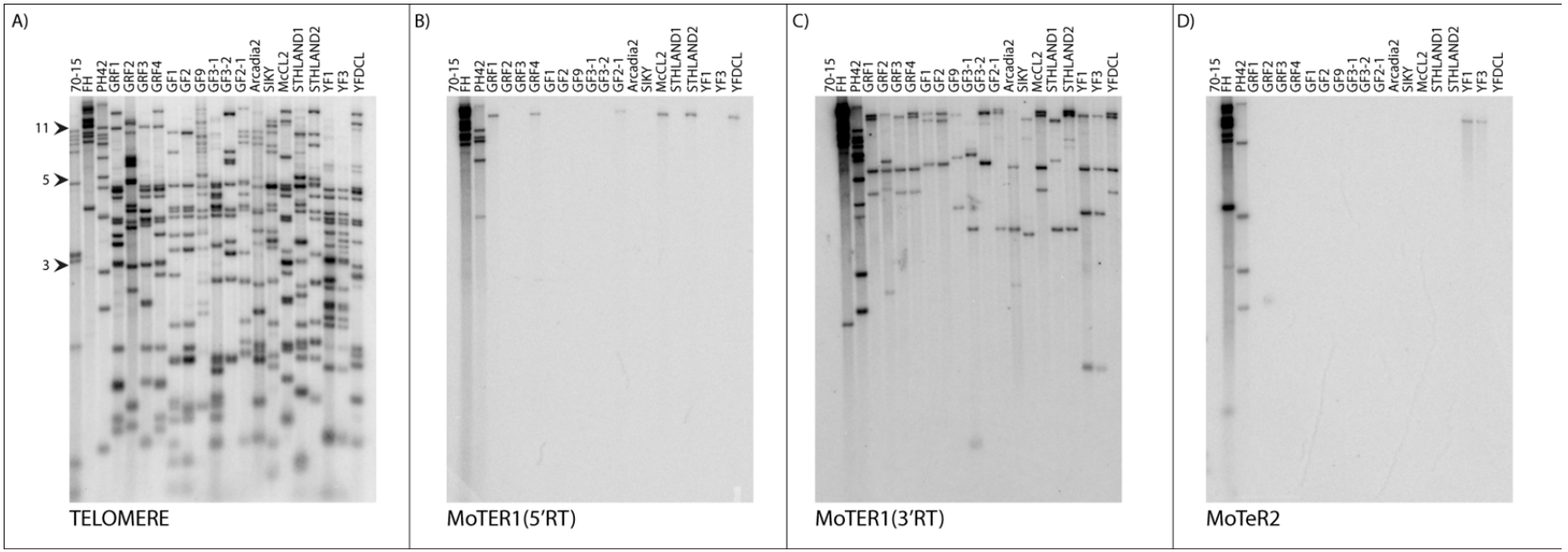

Figure 6-2.Southern analyses of genomic DNA from Setaria pathotype isolates of *M. oryzae*. Genomic DNA was digested with *Pst*I, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized sequentially with A) Telomere, B) MoTeR1(5'RT), C) MoTeR1(3'), and D) MoTeR2. The membrane was exposed to phosphorimaging screens and the resulting phosphorimages are shown. The membrane was stripped between probings. Lanes 1-3 are isolates used as controls 70-15 is the laboratory strain negative for the MoTeRs, FH was isolated from Lolium and is a positive control for MoTeRs, PH42 was isolated from Eleusine and is an additional isolate that is positive for MoTeRs. In lanes 4-21 are *M. oryzae* isolates that were collected from Setaria. Black arrows to the left of the blot indicate the size markers in kilobases.

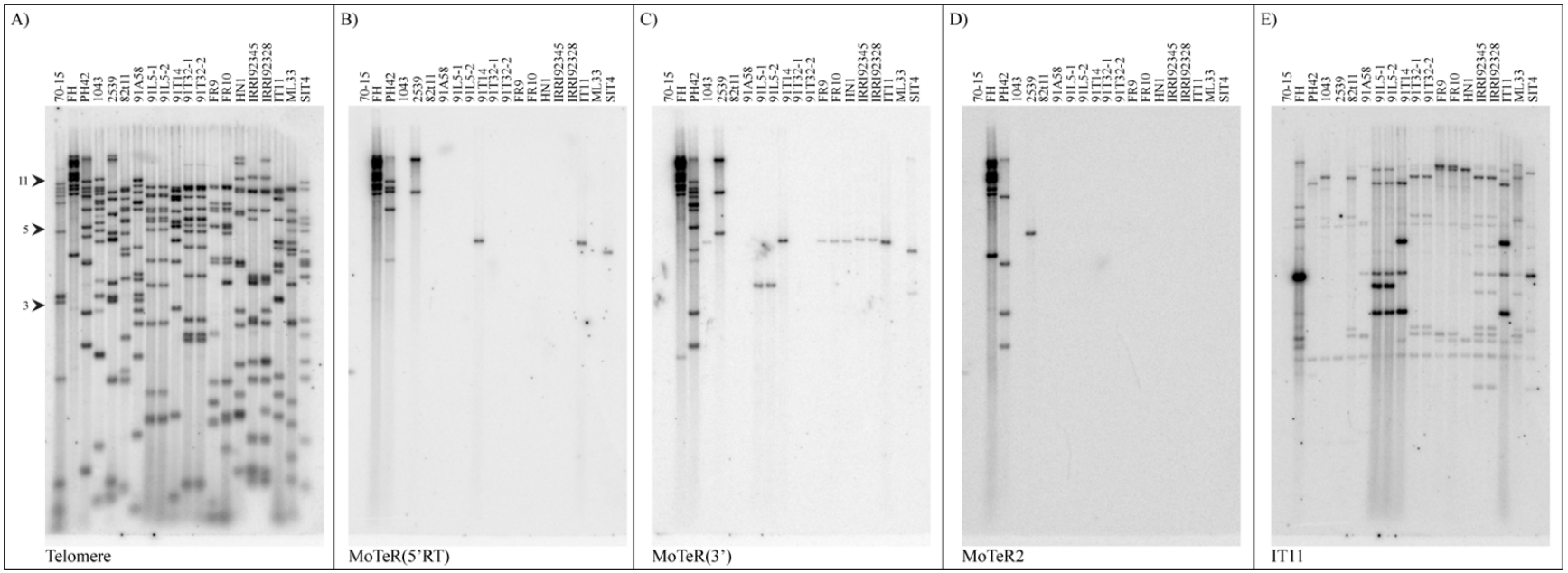

Figure 6-3. Southern analyses of genomic DNA from Oryza pathotype isolates of *M. oryzae*. Genomic DNA was digested with *Pst*I, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized sequentially with A) Telomere, B) MoTeR1(5'RT), C) MoTeR1(3'), D) MoTeR2, and E) IT11. The membrane was exposed to phosphorimaging screens and the resulting phosphorimages are shown. Membrane was stripped between probings. Lanes 1-3 are isolates used as controls 70-15 is the laboratory strain negative for the MoTeRs, FH was isolated from Lolium and is a positive control for MoTeRs, PH42 was isolated from Eleusine and is an additional isolate that is positive for MoTeRs. Lanes 4 and 6-21 are *M. oryzae* isolates that were collected from Oryzae. In Lane 5 sample 2539 is a lab created rice-infecting strain. Black arrows to the left of the blot indicate the size markers in kilobases.

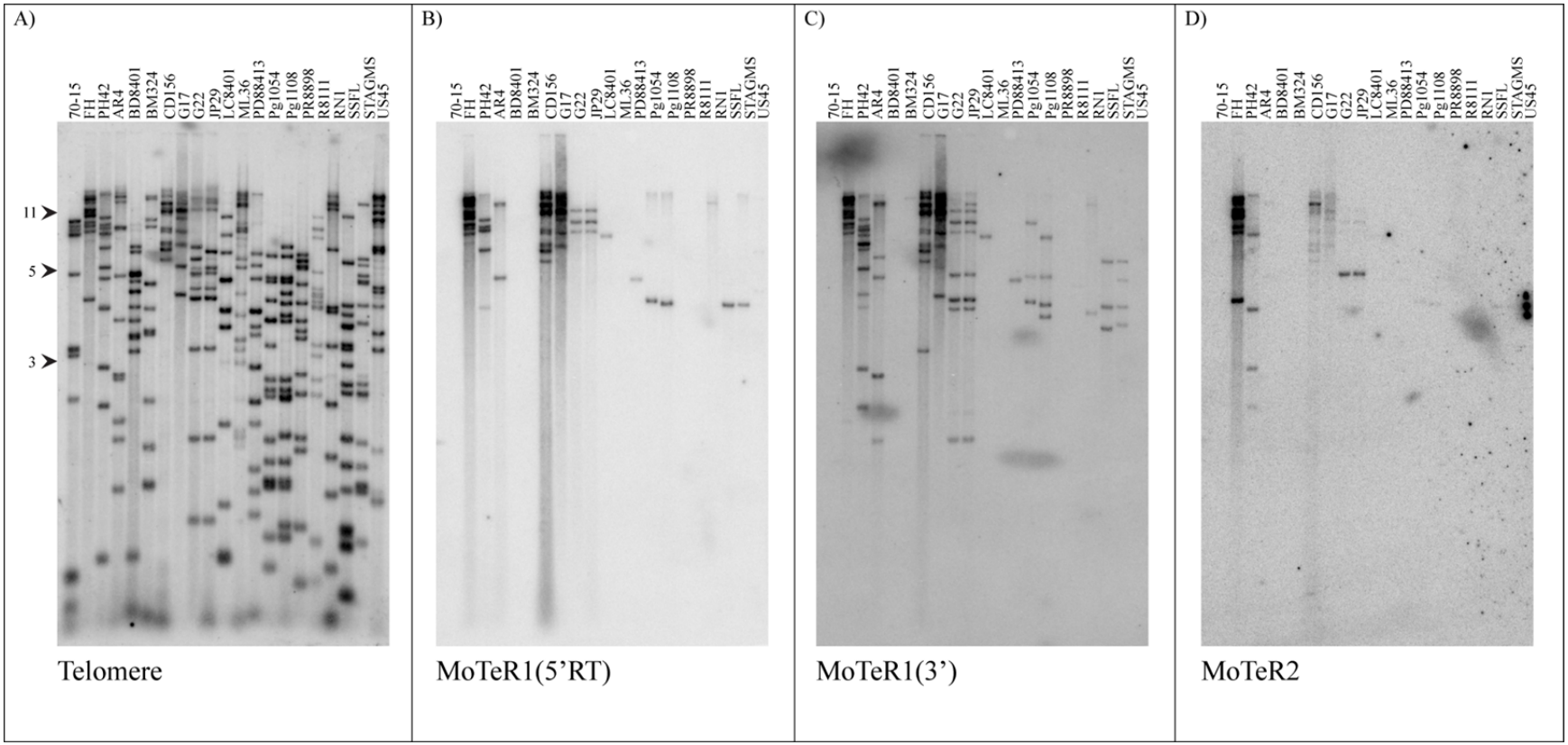

Figure 6-4. Southern analyses of genomic DNA from various pathotypes of Magnaporthe. Genomic DNA was digested with *Pst*I, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized sequentially with A) Telomere, B) MoTeR1(5'RT), C) MoTeR1(3'), and D) MoTeR2. The membrane was exposed to phosphorimaging screens and the resulting phosphorimages are shown. Membrane was stripped between probings. Lanes 1-3 are isolates used as controls 70-15 is the laboratory strain negative for the MoTeRs, FH was isolated from Lolium and is a positive control for MoTeRs, PH42 was isolated from Eleusine and is an additional isolate that is positive for MoTeRs. Black arrows to the left of the blot indicate the size markers in kilobases.

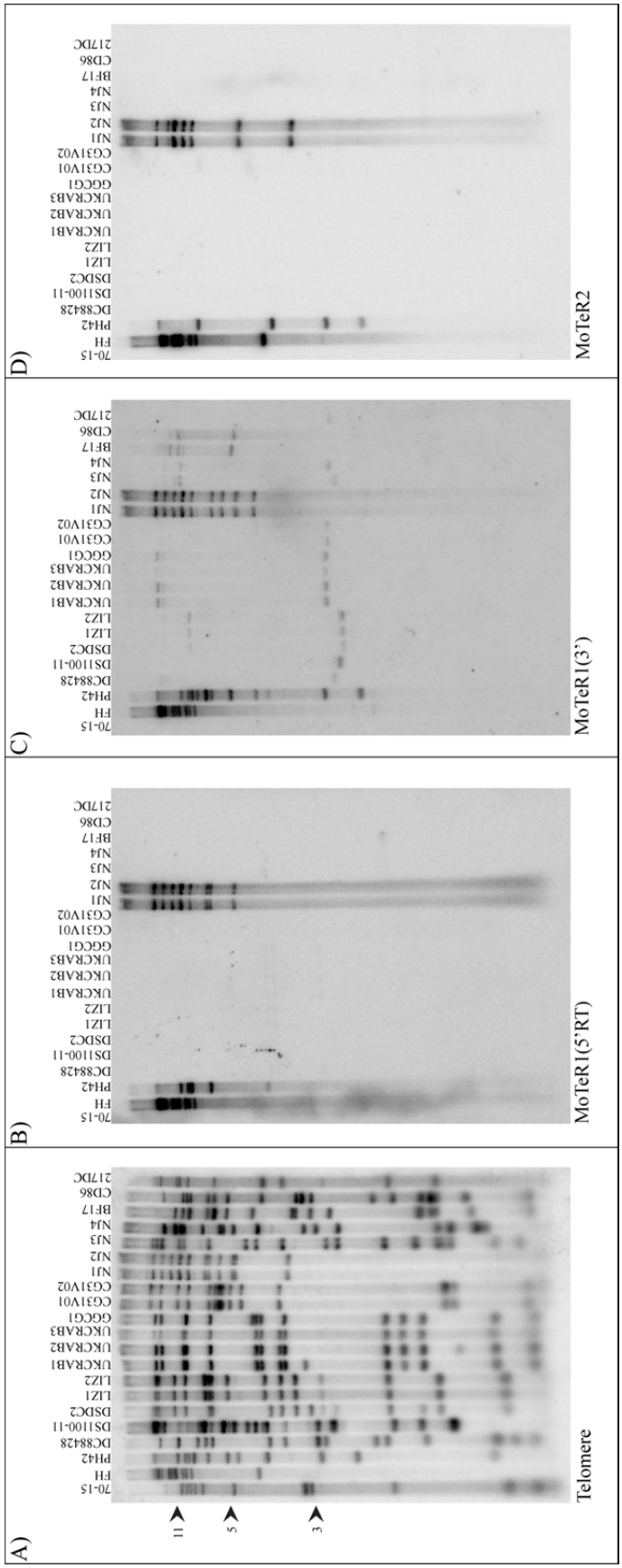

Figure 6-5. Southern analyses of genomic DNA from Digitaria pathotype isolates of *Magnaporthe grisea*. Genomic DNA was digested with *PstI*, electrophoresed in 0.7% agarose, electroblotted to a nylon membrane, and hybridized sequentially with A) Telomere, B) MoTeR1(5'RT), C) MoTeR1(3'), and D) MoTeR2. The membrane was exposed to phosphorimaging screens and the resulting phosphorimages are shown. Membrane was stripped between probings. Lanes 1-3 are isolates used as controls 70-15 is the laboratory strain negative for the MoTeRs, FH was isolated from Lolium and is a positive control for MoTeRs, PH42 was isolated from Eleusine and is an additional isolate that is positive for MoTeRs. The heritage of NJ1 and NJ2 (lanes 15 and 16) are currently unknown though it is suspected to be a Lolium-infecting isolate. Black arrows to the left of the blot indicate the size markers in kilobases.

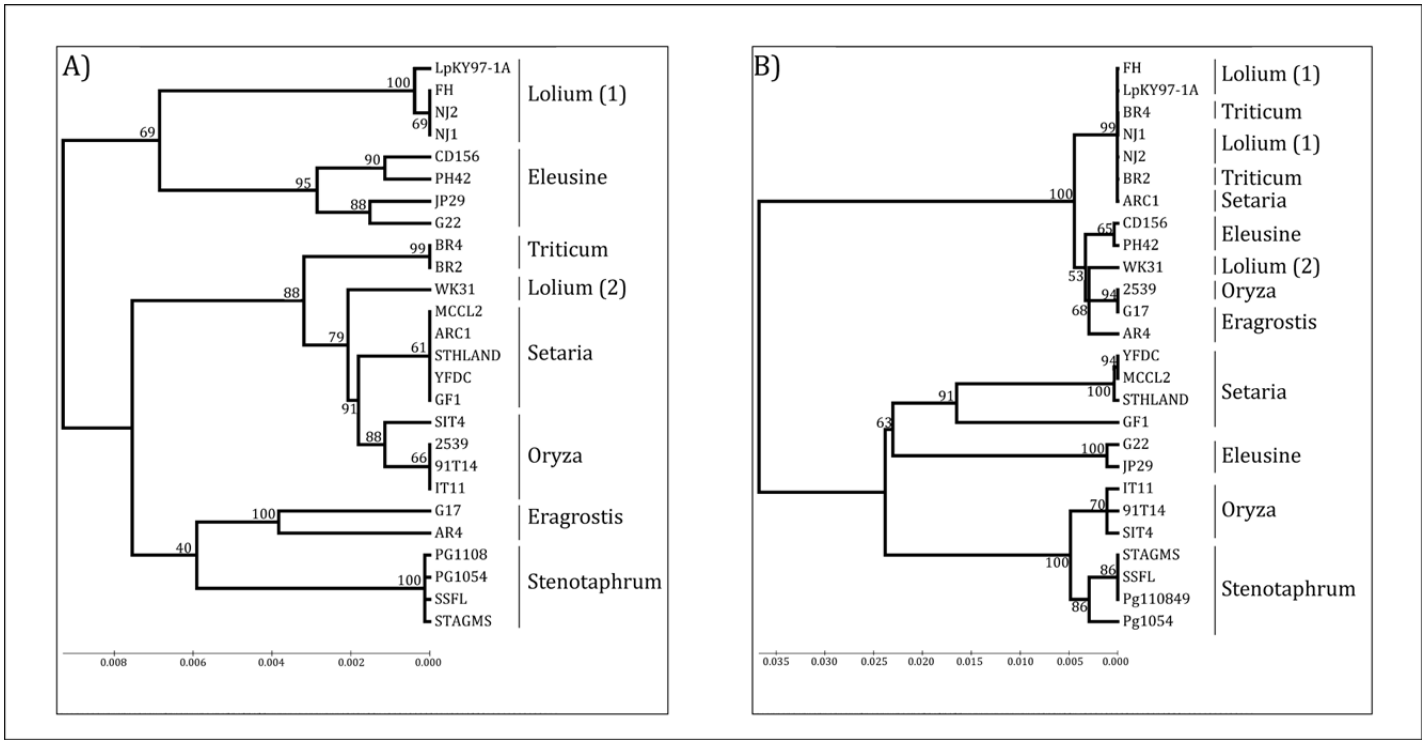

Figure 6-6. Neighbor-joining (NJ) phylogenetic trees based on concatenated sequence from three genomic loci [MPG1, CH7-Bac7, and the internal transcribed spacer (ITS)] (A) and MoTeR1 (B) in isolates from various *Magnaporthe oryzae* pathotypes. The bootstrap tree was inferred from 1000 replicates The percentages of replicate trees in which the taxa are clustered together in the bootstrap test are shown next to the branches. Missing data or alignment gaps were eliminated only in pairwise sequence comparisons. There were a total of 1323 positions in the gene loci dataset (A) and 1378 positions in the MoTeR1 dataset. The trees are drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. All analyses were conducted using MEGA version 4. Pathotypes are labeled to right side of the isolate name. Lolium (1) are isolates of the Lolium pathotype from the United States. Lolium (2) is an isolate was a weakly aggressive isolate that was collected from a lesion on perennial ryegrass (Lolium perenne) in Japan. Isolate 2539 is a lab created Oryzae pathotype strain.

CHAPTER SEVEN
Conclusions

7.1 MoTeRs are active retrotransposons in the isolates from the Lolium pathotype

In a previous study of GLS isolates, telomere profiles in Southern analyses were "hypervariable" in comparison with profiles of other molecular (FARMAN and KIM 2005). The major goal of this dissertation was to outline possible mechanisms that could account for this variability. By comparing mitotic progeny generated from single spored Lolium pathotype isolates and an Oryza pathotype isolate, it was shown that new telomere variability arises much more readily in the Lolium pathotype isolates. In fact, none of the third generation of mitotic progeny in LpKY97-1A had the same telomere restriction profile as the original starting culture. This suggested a molecular mechanism for the telomere profile changes. Southern analyses linked 70% of the telomere band changes in progeny of LpKY97-1A to MoTeRs. Because MoTeRs are embedded within the telomeres of LpKY97-1A it was believed that they could account for some of the telomere variation.

The MoTeR1 element has a variable 5' end, a 3' end, which is not adenylated post-transcriptionally, and an open reading frame coding for a reverse transcriptase with zinc finger, reverse transcriptase (RT), and REL-endo domains. Phylogenetic analyses of the RT domain indicated that this element was similar to other telomere specific non-LTR retrotransposons that were closely related to the site-specific CRE-like retrotransposons (Chapter 2).

Given that MoTeR1 had all the necessary components of non-LTR transposable elements, it was believed that they could be active within the telomeres of the Lolium pathotype isolates, where they were found in abundance. MoTeR2 did not encode a reverse transcriptase, but likely uses the RT machinery of MoTeR1 in its transposition. Preliminary expression studies indicated that both MoTeRs are expressed in LpKY97-1A (Appendix B). Additional results further suggested that MoTeRs were active within LpKY97-1A. Telomere-enriched shotgun

cloning approaches from different LpKY97-1A mitotic progeny found a probable *de novo* insertion of a truncated MoTeR1 in the telomere of the rDNA chromosome end and a possible transposition event of a MoTeR1 into a *de novo* telomere. Furthermore, expansions and contractions of MoTeR2 arrays were observed in mitotic progeny in Southern analyses. These examples lend support for the idea that MoTeRs are active retrotransposons in the isolate LpKY97-1A. The activity of MoTeRs could play a pivotal role in the instability observed in the telomere profiles in field isolates of GLS isolates. Further characterization of chromosome ends from mitotic progeny may lead to the discovery of many more such examples of MoTeR activity.

7.2 Possible role of MoTeRs in increasing the phenotypic variability

The instability of the telomeres in GLS isolates could have profound effects on the evolution and adaptation of clonal lines. MoTeRs are directly involved in the instability of the telomeres, but the phenotypic effects of these telomere profile changes have not been studied. One possible effect could be the change in expression levels of genes near the telomere end. This has been observed in Drosophila, whereby expansion and contraction of TART/HET-A arrays can affect the expression of neighboring genes (GOLUBOVSKY *et al.* 2001). Variegated expression patterns of cell surface genes allow parasitic organisms such as Plasmodium spp., trypanosomes, and pathogenic fungi to evade host immune responses (BARRY *et al.* 2003; VERSTREPEN and FINK 2009). The constant expansion and contraction of MoTeRs arrays could lead to progeny with differing expression profiles. One example where this could play a role in pathogenic variability is in altered expression of avirulence genes. If a gene was normally expressed, which conferred avirulence, then suppressing the gene product could lead to virulent progeny. Avirulence genes sometimes provide a positive fitness value to the individuals. By retaining the sequence in the genome, at a time when the fungus is no longer engaged with a R gene containing plant, the avirulence gene could be reactivated by moving to an area of the genome where it could be actively expressed. The 31B chromosome end in LpKY97-1A could be an excellent telomere to study more in depth. The readily observed expansion or contraction of MoTeR2 arrays in the 31B

telomere could lead to differential expression of genes in its subtelomere. Further cloning of that chromosome end would be needed to build a map of genes located on 31B chromosome end followed by expression analysis of those genes. This could be an interesting avenue to pursue in determining whether these expansions and contractions of the MoTeR2 arrays can modify the pathogen phenotype.

There are other possible mechanisms that could lead to diversity within the genomes of GLS isolates. Double stranded breaks could be caused by inefficient DNA replication through the telomere sequence between MoTeR-to-MoTeR junctions. This may lead to recombination with sister chromatids, truncated chromosome ends, or possibly strand invasion of the broken chromosome end into an internal chromosomal region followed by duplication of that sequence. Each of these mechanisms could serve to increase the genetic diversity of the mitotic progeny by shuffling the genome of an asexual line.

Truncations of the chromosome ends could potentially remove genes that could reduce fitness. The loss of a telomeric avirulence gene through chromosomal truncation, for example, could allow the progeny of a once avirulent strain to now cause disease in a host. This has been shown to occur in the Oryza pathotype isolates (CHUMA *et al.* 2011b). The accelerated telomere changes in the Lolium pathotype isolates could cause an increased loss of avirulence genes.

In at least one example an internal telomere sequence was copied onto the chromosome end in LpKY97-1A. If an internal gene is duplicated at the chromosome ends, modifications to the sequence could allow for novel adaptations. Other MoTeR containing telomeres could also serve as a potential homologous sequence with which to recombine through mitotic crossing over, which would allow for genes to be shared and even duplicated between different chromosome ends. The extra copy could provide for higher expression levels of the gene product. This has occurred in the evolution of other fungi including the genes involved in biofilm formation and carbohydrate utilization (CARLSON *et al.* 1985; MICHELS and NEEDLEMAN 1984; NAUMOV *et al.* 1990; VERSTREPEN and KLIS 2006). As more sequence becomes available from GLS isolates of *M. oryzae* it would be fascinating to look for such examples where

gene families have become expanded allowing for the exploitation of a particular niche in different isolates.

7.3 Implications of research to disease management

Mitotic progeny of GLS isolates showed instability in their telomere restriction fragment (trf) profiles and in their Pot2 restriction profiles. The trf profiles showed a much higher instability than the Pot2 profiles. Since a majority of avirulence genes are located near telomeres or near transposable element clusters (FARMAN 2007), instability of this region brings into question the efficacy of employing new resistant plant cultivars that are based on using a single R-gene. Even the more effective approach of deploying multiple R genes in new plant cultivars may encounter problems with virulent strains of the GLS pathogen emerging readily. A better approach to control the GLS pathogen would be based on quantitative disease resistance, which tends to be more durable than R-gene mediated resistance (PARLEVLIET 2002). Resistance breakdown is less of a problem in quantitative disease resistance because of lower selection pressure on the pathogen due to the smaller effects of multiple genes. This is due to the somewhat inconsistent effects and only partial resistance of multiple quantitative resistant genes in which the pathogenic variants that overcome this type of resistance only gain slight advantages. In R-gene mediated resistance there is a stronger selection for pathogenic variants (POLAND *et al.* 2008). Using quantitatively resistant cultivars with other management strategies that limit the fitness of the pathogen could curb the number of severe epidemics in turfgrasses.

The high mutability of isolates with active MoTeRs could lead to host switching if an avirulence gene was lost in the truncation of the chromosome end. There is some evidence to suggest an invasion of a non-Lolium pathotype isolate may have occurred in the WK isolates, a weak pathogen of perennial ryegrass in Japan (TOSA *et al.* 2004). This could be an issue if recombination or horizontal transfer occurs between the Lolium pathotype isolates and the newly invading isolate due to this new close association. Fungicide resistance and virulence have been transferred by parasexual recombination between isolates of *M. oryzae* in the laboratory (NOGUCHI *et al.* 2006), which adds another level of complexity to dealing

with gray leaf spot disease over the long term. Additionally, the horizontal transfers of supernumerary chromosomes have been suggested to occur (AKAGI *et al.* 2009; HE *et al.* 1998; MA *et al.* 2010). Supernumerary chromosomes are extra chromosomes containing DNA that may not be found in all representatives of a species (COVERT 1998), and in other pathogenic fungi they can play important roles in pathogenicity (HAN *et al.* 2001; HATTA *et al.* 2002; MIAO *et al.* 1991). With other species of grasses commonly bordering managed perennial ryegrass fields the fungicide resistance developed in another host pathotype of *M. oryzae* could be transferred into the Lolium pathotype isolate. Thus a turf manager may need to be attentive to possible invasions of *M. oryzae* from other hosts. Detection of new invaders into perennial ryegrass may be important as plant breeders may not take the potential aggressiveness of other host specialized isolates into account when developing resistance in plant breeding programs. The extreme variability of telomeres in GLS isolates and the constant generation of new variant progeny will lead to continued challenges in controlling GLS in perennial ryegrass. This dissertation has addressed some of the mechanisms involved in the generation of the telomere variability, but work is still needed to understand the implications of this variability in disease management.

Appendix A

Sequences of MoTeR1, MoTeR2, and NhTeR1

Magnaporthe oryzae Telomeric Retrotransposons sequences

>MoTeR1
GAACCCGAACCCAAACCCAAACCCAAACCCAAACCCAAACCCAAACCCAAACCCAAACCC
GGAGGGTTCCCAAGTCGCCTAAACCCGAAGGGTTTAGGATATTATTTCGTTTATTAGAA
TTGGATAATTATTTACCCCTGTTGGACAGGGGGGTTGCAGGGGTTAAATTAAGGTTTTT
TATTATTTATGCGCCGTTTATTTGTTTACCCCCCCAAATATTATAAAAGCGCGTTCCATC
CTCTTAGGAAAAGCGAAGCTTTTCCTTGTAAAAGTCGCTAGACTTTTACTATAAAAGTC
GCTAGACTTTTATACCAATCTTTTAACAAAAAGCGTAGCTTTTTGTTGCCAATCTATTAA
AAAAAGCGGAGCTTTTTTTAACTTTTTCTTTTTTTTTTTTTTTTCTTTTTTTTTTTTTTT
TTTCTTTTTTTTTTTTTTTTTTTTTTTATATATATTATTATTATTATTATTAGCGGTGGGG
CTATTTATGCGCTTTAATTTGTGCGGGGCTATTTATGCGCTTTAATTTGTGCGGGGCTAT
TAATGCGCTTTAACTTTACAAATTTTATTTATGCGCTTTAATTGCTGCGGGCCTGTTAAT
GCGCTTTAATTTACAAATTTCATTAATGCGCTTTAACTTTTATATTTACTAATGCGTTAT
TTATATAATTGCTATTATTATCGTTGCTATTATTATTATTGCTATTATTATCGTTATTAT
TATTGCAATTTTATTATATAAACCCTCGTTTGTCCCTCGATTTATCCCGTTTCTTTTCCA
TCCCATCGCGCGTTTTCGTAAGCTTTGGTTTTCGTAGGATTTGCTTTCGTAGGCTTTGCT
TTCGTAGGCTTTCGTCAGCTTTTACCTGCTTTTATTTTTTCTTTTTCTTTTTATTCCCCCC
CCTTTTTTTTACCTGGTTTATTAGCGGTTTACCTGCTTTTATTACCTGGTTCCCCTTTAC
CTGTTTTATTAGCGGTTTACCTGCTTTTATTACCTGGTTCCCCTTTACCTACTTTATAAG
CGGTTTACCTGCTTTTATTACCTGGTTCCCCTTTACCTGTTTTATTAGCGGTTTACCTGC
TTTTATTACCTGGTTCCCCTTTACCTGTTTTATTAGCGGTTTACCAGCTTTTATTACCTG
GTTCCCCTTTACCTACTTTATTAGCGGTTTACCCGTTTCTATTAGTGGGCATTTATTTCC
CGTTTTTATTAGCAGTTAAATTTACCCTTTTAAGGTTATTTACCTGCTTTTATTCACAGG
GCACCCCTGTTTTTACTAGCAGTTAAATTTACCTTTTTAAGGTTATTTACCTGCTTTTAT
TCACAGGGCACCCCTGTTTTTACCAGCAGTTAAATTTACCTTTTTAAGGTTATTTACCTG
CTTTTATTAACAACCCTTTATTTTTTCCTATTAACGGGTATTTATTTACCTGTTTTATTG
GAATTCACCCGTTGGACGGCATGGTTTGCCCAACCTGTAACGGCGTTTACGCCGATTACA
ACGACCATATCCGGAAAAAACACCCGGACGAACGTTATACCGCCCTCCAACTCCAACCAT
TGGGTTTAACCCCCTGCCCTATATGCAAAACCGCTTGCAAAAACGATTTGGGCGTTAAAA
CCCACCTATCCAAAATCCACAAAATATCCGGTGCATCGAAAATTTCAACCCAACCGCGTA
TACGAACGGAAAATACGGATAATACCAATTCGGTCCCCACGTCGTCGTTTAACCCTGTCC
TTCCCGAAATCCAAACGTTAACCCCGGGGTTAAATAACAGCCGTTGGGCCGATAACCCCA
GAAAACGACGGGCCGATACCCCCTCCCCAACACGGGGTCGGAATACACGCCCACGTCGAT
TTTCATATACGGATATCGATTTAACAAACGACGAACCGGCGGATAACCCCAGGGCTAATA
ACCCCAGGGTTAATAACCCCAGGGTTAATAACGAACCCCCCTCCAGCCCAAATTCGTTAC
CTTCGATTTCCGAATTTCACACCCCTGGGACCCTACCCCTAACCAATTCGAATATATCGT
TAAAAGACCAGCACGACAAAATTACCGGCCCTATATTGCAAAAACCGTTAATCCAAAAA
TTAATCGAATATTCGAAAATCCCAATCCCAGAACACCACCTCCACGCCAGGCAGGCTAAA
ATTTTTGCTGACGCCGCAAATCGAATCGCCAAAAATTTTATACAAAGCCCAACGGAGAA
AACATTATTTAATTTACTTATATTACCCCGCATATTCGGTATCGGGTTAATAAACGGAA
AAGTAACTAAAATAATGCAAAACTTCCCATCCCAAATACCCCCTATTCCAAAAATTGATT

TTCCATCCGAAAAAACCGATTCCGACCCGGTTTTAAACGCCAAAAAATTATTGGAAAAA
GGGTATATTGGCCGTGCGGCAAAGGCTATTATCGATCCAACCCCCGTTGCCCCAGAAACC
CCGGAATCGTTAAATATTTTACGGGAAAAACACCCTATTGGCCAAAATAACCCGTTTAA
TACAAAATCCCAACCAATATCAGGCAGGCAAATTACCGAAAAAGCTATTTTATTAGCTA
TTTCGTCTATTGGCCGGGAAAAAGCTCCGGGCCTTAGCGGGTGGACGAGATCGTTATTAG
ATGCAGCCATTAAAATACCTACCCAAAACGACGTAATTCCGGCTTTACGACTCTTAACGG
ATATGATTCGCCAGGGTACCGCACCGGGTAGGGAATTATTATGCGCTTCGCGTTTAATAG
GGCTATCCAAACCCGACGGCGGCGTACGCCCAATAGCCGTTGGGGACCTATTATATAAAA
TAGCCTTTAAAGCTATTTTAAATACCCTATGGTCCCCAAACTGTTTATTACCTTACCAAT
TAGGTGTAAATAGTATAGGTGGCGTCGAACCCGCTATTTTTACCCTCGAAGAGGCTATA
ATGGGCCCTAATATTAACGGTATAAAATCGATTACCTCCCTCGATTTAAAAAACGCGTTT
AATAGCGTATCCAGGGCTGCAATAGCCTCGTCGGTAGCTAAATACGCACCAACTTTCTAC
CGTTCTACCTGTTGGGCCTATAACCAACCTTCGATTTTAATAACGGAAAACGGTTCCGTC
CTGGCTAGTGCACAAGGTATACGCCAAGGCGATCCGTTAGGCCCGTTGTTATTCAGCCTT
GCTTTTCGACCTACGTTGGAAACGATCCAAAAATCGCTTCCATATACGTATATAGCGGCT
TATTTGGACGACGTTTATATTTTATCCAAAACGCCCGTTAAAGATAAAATAGCCAAAAT
AATCGAAAAAAGCCCGTTTACCCTAAATTCCGCCAAAACGACAGAAACGGATATCGATA
CGTTAAAAACCAATGGTTTAAAAACGCTCGGCTCGTTTATTGGACCAACGGAATTACGG
AAGGAATTTTTGCAAAATAAAATTCAAAATTTCGAATCGTCCATTAACGCCCTGAAAAA
ACTCCCTAAACAATACGGATTGCTAATCTTGCGTAAAAGTACACAATTACTTTTACGCCA
TTTGCTCCGTACTTTAAATTCCCAGGACCTGTGGGAATTATGGGAAAAAACAGATAAAT
TAATAGCGGATTTCGTTATAAATTTAACTGTTACAAAACGGAAAAAACGGCCAATTACG
GATTTCGTTACGCCGTTAATTACGTTACCTATAAAGGACGGAGGTTTTGGATTATTACG
GCATAACGGAATAGCCCAAGATATTTATTTTGCGGCCAAGGATTTAACAACCGAAATTC
GGCACAAAATCCAACGTATATCCAACGATTTTCCACAAAATCAAAGCCCTACCGCCACCG
AGATTTTGCATTTGTTGCATAACGGGGTTTTAGCAGATTGCAAAAACGGGTTAACAAAC
GCCCAATTAAACGCTTTAACCGAAAACGCTAGTTATTTAGGTCGAAAATGGCTTAACAT
TTTACCTATCCAAAAATCAAATCGATTAACGGATTGGGAAATGGCTGAAGCCGTTCGAT
TAAGATTATTAGCCCCGGTTAAACCGTTAACCCACCCCTGCAACCATTGCGGAAATCGGA
CCAATATAAACCACGAGGACGTTTGCAAAGGTGCCGTACGCAAATATACGGCCCGTCACG
ACCAAATAAACAGAAGTTTCGTCAATTCGTTAAAAAGTCGACCAGAAATCGACGTCGAA
ATCGAACCCGATTTAAATAACGAAAATAACGTAAATAACGCCAATACAACCACCGAAAA
TCCCACCCCTAGCCCCAACGGCCAAAACGATACCGGATGCCTTTTTACAACCCCTATTCGC
TCCGGGACCCGTAACGGCCAAAACGGCCTTAGGGCGGATTTTGCCGTTATTAACGGCGTA
TCCAAATATTATTACGACGTGCAAATCGTTGCAATTAATAAGGATTCCGGTAATACAAA
TCCGTTAAATACGTTAGCAGACGCAGCAAATAACAAACGACGTAAATACCAATTTTTGG
ATCCATTTTTCCATCCAATTATAATAAGCGCCGGAGGCCTTATGGAAAAGGATACAGCAC
AGGCGTACAAACAAATCCAAAAATTAATAGGCCCCGTTGCGGCCCATTGGTTGGATACGT
CGATTTCGTTAATTTTGTTACGGTCCAGAACGACGGCAGCAATTTCTATTGCTAAAAACC
GCCCTCGTGCGTAATAGGTAACGTCCCTATTTTTGTCTTTGGTTTTGTTTTTATCTTTGT
TTTTGTTTTTGTTTTCGTTTTTGTTTTTGTTTTCGTTTTTGTTTTTTTTTTTGTTTTTGT
TTTTGTTTTTGCCTTTGTTTTTGTTTTTATCTTTATTTTTGTTTTTGTTTTTACTTTGTT
TTATTTGTTTTATATTTACCTTTTGATTTTTTCTATTTTTCCCACCCTTATTATTATAAC
CCCAACCTACTAATATTTTTTCTTTTTTCTTTTTTCTTTTTACGGTTTTATTTTCCCGTT
TGTTTTTTCTATTTTATTTGTACGACAAAACCCTTAGCAAATAAGCTTAGAATATAATA
AAGCGCGAATTAAAACCCTAACCCTAA

>MoTeR2
GAACCCAAACCCAAACCCAAACCCAAACCCAAACCCAAACCCAAACCCAAACCCAAACCC
GAAGGGTTCCCAAGTCGCCTAAACCCGAAGGGTTTAGGATATTATTTCGTTTATTAGAA
TTGGATAATTATTTACCCCTGTTGGACAGGGGGGTTGCAGGGGTTAAATTAAGGTTTTT
TATTATTTATGCGCCGTTTATTTGTTTACCCCCCCAAATATTATAAAAGCGCGTTCCATC
CTCTTAGGAAAAGCGAAGCTTTTCCTTGTAAAAGTCGCTAGACTTTTACTATAAAAGTC
GCTAGACTTTTATACCAATCTTTTAACAAAAAGCGTAGCTTTTTGTTGCCAATCTATTAA
AAAAAGCGGAGCTTTTTTTAACTTTTTCTTTTTTTTTTTTTTTTCTTTTTTTTTTTTTTT
TTTCTTTTTTTTTTTTTTTTTTTTTTTATATATATTATTATTATTATTATTAGCGGTGGGG
CTATTTATGCGCTTTAATTTGTGCGGGGCTATTTATGCGCTTTAATTTGTGCGGGGCTAT
TAATGCGCTTTAACTTTACAAATTTTATTTATGCGCTTTAATTGCTGCGGGCCTGTTAAT
GCGCTTTAATTTACAAATTTCATTAATGCGCTTTAACTTTTATATTTACTAATGCGTTAT
TTATATAATTGCTATTATTATCGTTGCTATTATTATTATTGCTATTATTATCGTTATTAT
TATTGCAATTTTATTATATAAACCCTCGTTTGTCCCTCGATTTATCCCGTTTCTTTTCCA
TCCCATCGCGCGTTTTCGTAAGCTTTGGTTTTCGTAGGATTTGCTTTCGTAGGCTTTGCT
TTCGTAGGCTTTCGTCAGCTTTTACCTGCTTTTTTACCTGCTTTTATCACTTGTTTTTAT
TTCCCTTTTACTTTCCCTTTACCTGTTTCACAGGTATTTATTATGGATTTATTTATAAAC
CCCCCAAACCCAACCCCCGACCTCGACCCCGACCCTGATCCGGACCCCGACCCTGACCCGG
ATCCCTATCCAAATATTAACGCCGCCGTCGATTCGTCCCGCCAAAAATCAAATATATATA
TCGATTTAAATTCCAAATTTAATTCGGTTAACCCCCGGTATATTAAAGTCGCTAAAAAA
TCCTGGAATATACGTGCCTTTTTAAAACAACTTTTTGCCGTCCCTATCCAGATAACATGG
TTTTTTAGCAATGTTATTATCCACGGGTTTACCAATTCTATATTTGGTATTTATTCGATT
TATTTATTCGATTTTAACCCCCGATTTCGACCGACTATTATCGATTTATTACGCCAAAAG
TCCAGCAAATATACCGATTTAAATCCCGAATTTGAATTGGCTAACCCCCTGCATATTAAA
TTGGCTGAAAAATCCTGGAATATACGTGCCTTTCTAAAACAACTTTTTGCCGTCCCTATC
CAGATAACATGGTTTTTTAGCAATATTTCTATCCTCAGGGTTAACCAATTTTATATTTGG
TATTTATTCGATTTATTTATTGGATTTAATCCGCTAAATAAATTGTTAATCCGTTAATT
ATATTGGAATTAATCCCCGAAAGAGGACCAACCCAATATATATTCCGATAGGGAATTTT
TTCTTTTTCTTTTTTCTTTTTTTTTACGGGTTTATTTTTTCTACCCTATTTGTACGACAAA
ACCCTTAGCAAATAAGCTTAGAATATAATAAAGCGCGAATTAAAACCCTAACCCTAA

Nectria haematococca Telomeric Retrotransposon sequence
>NhTeR1
TTAGAACAGAGTAACTATTAACTATCTAAGCTATTTTGCCCTAGGGGTTTCTTCTAGCTA
AGATATAGCTTATAGATAAAAGAGGGGAGATTTTTCTAAGAAATAATAAAAGAAAAAG
CATAAGAGATATAACTAGGAAAGAGGTATAGCTAAGAAAGAGATAAAGCTAGGGGAGA
GATATAGCTAGGAAAAAGAGAAATAACTAAAAGAGAGAGGTATAGCTAGTAAGAAAAA
GGCAAAGAGATAAGAGATAAGAGATAAAAGGTAAAGGATAAAGAAAAGCCTTAAGAAG
CCCTAGGGGTATTCCAGATCGACCTAGCTAGGTTACTAGGGGTATCTCGGGCTATAGAAG
CCGCAGAGGTCGCCCTAGTCCTCATCAGGGCAAGGCTAATAGAGGAATCTAGCTGGTTAG
CCGCAAGGGGGCCAATAAGGTCCTGGAGCTTCTGATAGGTCTTAGCTGTCTCTAGGTCCA
TAAGGCCGCCTGCTGAGATAATAAGAGGCTGAAAAAAGGCCCCGAGAGACCTATATTTT
CTGCGTTTTTCTTCGGCAGCCTCTCTTAAGGTGCTATAAGGGTCTTCTTTAGCAGAATCC
TTTGAAATAGCTACGATTTGAATATCATAGTAATAGCGGCTATTTCCTAGGTTAACTAC
AAAGTCGGCCCTTAGGCTACCTTCTGGGTGGACTAGGGGTTCTTTTTCGACCTCTAGGGT
AGGCTGGCTACTTAGGGCTCTCACAAAGGCCCTAGTTATAGCGTCATGCCTTGCAATCCA

CCTTCTATTAGCACCCTTGCAGGTATCCTGATGGCCTAAGGCAGCTATAGCTCCACAAGA
GCTACAGGGTAAGCTGATAGGCTTGATTGGGTAGAAAAGCCGGCTTCGGAGGGCTTCAG
TGACCTCAGAGTCTGTAAAAGATAGGGCTTTTTGGGTAGGTAAGACTCCGAGCCACTTAC
GGCCGAGATAGCTAGCATTTTCTAGCCTAGCATGCCTATAGCTAGCGGGGAGGTCCTCTA
AGAAGCCTGCTAGCCTAGCCTTATTAGCTTCCTTTAGCACCTGTTGGGCTGTTTTCCCTA
GTCTAGGCTGGCCTTGGTTAGGCTGGCTTTGGTCTAGGGTAGGCTGGGCTAGGGGCTTCC
GGATAAGGCCTAGGGTAGGCTGCGAGGCCTCCCTAGCTGCTAGGAATAGCTCGTGGGCTA
GGTCCTTGTGTAAGGGTATTCCTAGGCCTCCCTCTCTGACTGGAAGGGCTATAAGGCTTG
AATTAGGCTCTTTAGGGCTCTCACTAGGGCTTCTAGCCACTAGGGCTATAATAGCCTCTC
TTATAAGGGTATCAGCCTCTTCCCAAAGGTCTTCTAGCCCGGTTGGGTCTAGCTGCCTCT
GGAGGTGCCTTAGAAGAAGCTGGATACTGCCTCTAAGAAGAAGCAAAGAGTGCTGCTTT
GGAAGGTCTTGCAGGGCCTCTAGGGCTTCCTGTAGGGTAGCTAGCTTTTCTTGCAAGAAG
GTTCTTTTAGGCCGGATAGGGCCGATATAAGTACCTAAGGCTTTTAAGCCTTCTAGCTTA
AGGTCCTCTATAGCCTTTTCCTTGCTTTTAGCTAGGTTAAGGCTAAAAGGGGAGCCTTTT
AAGACCTCCTTAGCTGCCTCTAAGGTGCCTTGCGGAGCCTTATTAAGTATATAGAGGTCA
TCTAGGTAAGCTACTAAGGTTGCCCTAGGTAGCTTCCTAGCTAGGGCCTCTAGGGTAGGT
CGAAAGGCAAGGGAAAAAAGAAGAGGTCCGAGAGGGTCACCTTGCCTTACGCCCTTGGC
CGAGGCTATAGCTGACCCATCTTCCATAACTAAGATAGAAGGGTCATTATAGGCCCAAGC
TGCTGCCTTATAGAAGGTAGGGGCAAAAGTAGCTACAGAAGCAGCTATAGAAGCCCTAT
CTACTGAGTTAAAGGCATTAGCAAGGTCTATAGAGGCTAGCTGCTGAAAATTAGCCTCA
TTTAAGCCTATAATAGCCTCATAGAGGAGGAAAATAGCGGGTTCAACCCCGCCTGGGCTA
TTCACGCCTAGCTGGAAGGGTAAGAGCATATTCGGCCGATAGGAGGTCATTAGGATCGCC
TTCATAGCTACCCTATAGATAAGATCTCCTATAGCTATAGGTCTAACCCCTCCATCGGGC
TTTTCAAGCCCTATAAGGCGGCTAGCGCATAATAGGTGGGCTCCAGGGGCCGTGCCTTGG
CGAATCATATCAGCTAGCAGCCTTAGGAAAGCTATCACAGGGGAATCTTTCCTAGTTACT
AGGTCTAGAAGAGGCCTAGTCCAGCCACTAAGGCCCGGGGCCTTTTCCTTGCCTATAGAG
GCTATAGCTGCTATAATAGTCTCTGACGTGATTGGTTGGCCGGCCCTTGGGCGGGTCTTG
CCTTGGAAGGGGTCTTTTGATCCAATAGGGTGCTTTTCTAGCAGCCTAGCCCTATTTTCT
ACTGAATTAGGTGCTATAGGGGTCGGATCAATAAGAGCCCTAGCAGCCCGGCCTAGGTA
GCCTCTCTCTAGTAGCTTAGCTGCCCTTTGGGCCGGGCTAGGAGCTATAGAGGGTCTAGC
TGCCTTGGGGGGCTCAGGGGGCTGCTGGAGGGACTCTAGGCTCTCTAGGGTAGGTAGGTT
AGAAGGGAAAGACCTAAGAAGGGTAGCTAGGCCTCCTTTTTGTAACCCTAACCCTAGTA
GCCTAGGCAAGATAAGGAAATATAAGAGGGCCTTTTCAGTAGGCCTCTTTAGAAAGGCA
GCTGCAGCTTTATGGGCTGTAGCTGTGAAAATAGCAGCCTGCCTTGCATGTAGCCTTTTT
TCCGGTATAGGGATCCTTGCAAAGGCTAATAGCTTCTGCATAGAAGCCTTGGCTAGGATT
GGGGCTATAGCTTGGTTGTGTAGCTGCTCTAGGGTAGGCCCCTCTAGGGTAGGCTCCTTG
GGCTCGGGCTCTAGGGTAAGCTCCCTAGGCTCGGGCTCTAAGGTAAGCTCCCTGGGCTCC
CTAGGCTCCCTAGTCTCTAGCTCTAGGGTAGGCTCCTCTAGGGGCCCTAGAGGCTCTAGG
GTATATTCTTCTAGAGGCTCTAGGGTAGGTCCCTGACTAGAGGGCCTGTGGCTAGGCGTG
CTAGGGCAAGAGGAGCTAGAGCTAAGGGACCTAGGGCTAGAGGGTCTTCCCCTTTCCTCA
TAGCTAGAAAAAGACCTATAGCTTTCCCTTGAGGCCGGTCGGTCTAGGGTAGGCTCCTCT
AGGGTAGAGTTAGGCTGTAGCTGCCTTTGCCGCCTAGTAGGTCTTTGTAGCTGTGGAGAA
GGGGTCCTTGCTAGTCTCTTTCTTCCTATAAGGCCTCTATAGCTAACTAAGCCCGGGCTA
GCCTGATTAGGACTAGTGGGGCTAAGGGAAGAAGCTAAAGAGGCCCTAGAGGTATTAGT
GGCCCTAGTAGCTATATAGGTCCTTGGCCTAGGTAAGGTAGAAATATGGGCTTTACCTTC
TATGCCGTGGATCTTAGCACTATGGGTCTTGATGCCGTGACTTCCGCGGCAGGCTGTGCC

GCAATAAGGGCAGCTGACTAAGCCAAGGGGCTGTAGCTGTTGGTTGGTATAGGCCTCGGC
CGGGTGTTTTTTGCGTATATGCTCTAGGGTATCTCTATAGCTGCCTTGGCAGGTAGGGCA
TACCTGCGAGGGCATTTTACTATAGGGTAGCTAGAGGTATAGGGTAGGTATAGGGTAGC
CGAGGTATAGGGTAGGTATAAGGTAACTAGAGGTATAAGGTAGCTATAAGAAAGAAGAT
AATAAAGAGAAAATAGCTATAAGGTAGCTAAGATAAGGCTTAGAAACTAAGCATAAGGG
TAAGAGTTAAGGTTATAAGAAGCTTAGAGAATAAGGTCTTTAGAGATACTTATATAACC
CTAGGGTTATTCTAGAAGCTTCTACCTTAACCTAAGTCTAACCCTAAGCCTAACCCAAAG
CCTGACCCTAACCCTGACCCAATGTGTGACCCTAACCCCTATAGAAGGCTCTAGAATCTC
TATAGAAGCCCGAAGGGCTTCTATCTATAGCTGACCCAAAGCCTGACCCTAACCCTGACC
CAATGCCTGACCCTAACCATGACCCTAACCTCTATAGAAGACCTTAGAATCTCTATAGAA
GCCCGAAGGGCTTCTATCTATAGCTGACCCAAAGCCTGACCCTAACCCTGACCCAATGCC
TGACCCTAACCATGACCCTAACCTCTATAGAAGACCTTAGAATCTCTATAGAAGCCCGAA
GGGCTTCTATCTATAGCTGACCCAATGCCTGACCCTAACCTCTAGAACCTCTAACGGAAC
CTCTAAAGGAACCTCTAGGGGGATAGGGTGAGGGATAGGGTGAAGGATAGGGTGAGGCA
TCGGG

Appendix B

MoTeR1 Expression Preliminary Experiment

B.1 Introduction

Active Non-LTR retrotransposons transcribe an RNA intermediate in their transposition. *Magnaporthe oryzae* Telomeric Retrotransposons (MoTeRs) were recently discovered in the telomeres of gray leaf spot (GLS) isolates of *Magnaporthe oryzae*. There are two different MoTeR retrotransposons. MoTeR1 is a ~5.0 kb element that has an open reading frame for a reverse transcriptase, while MoTeR2 is a shorter ~1.7 kb element that has an open reading frame for a protein with no known function. To determine if MoTeRs were actively being expressed reverse transcriptase PCR experiments from total RNA were used.

B.2 Methods

Mycelium was grown at room temperature with shaking for 7 d. The mycelium ball removed from the 10 ml CM using sterilized forceps, and blot dried on paper towels. Total RNA was then extracted from the mycelia cultures of LpKY97-1A using the protocols outlined in the RNeasy Universal Mini Kit (Qiagen, Valencia, CA). Total RNA was double digested with DNase to remove DNA contamination using the protocols outlined in the Turbo DNA-free kit (Invitrogen, Carlsbad, CA). RNA was incubated at 65°C for 15 minutes and then chilled on ice for 5 min. First-strand synthesis reactions were completed using the SuperScriptII First Strand Synthesis for RT-PCR kit (Invitrogen, Carlsbad, CA) with 2 μM of MoTeRstartR (5'TTTTAATTCGCGCTTTATTA3') or a 50 μM of Oligo$(dt)_{20}$ primer. Following first strand synthesis the samples were treated with 1U of the restriction

enzyme *Mbo*I (New England Biolabs, Beverly, MA) at 37°C for 30 min followed by heat inactivation of the enzyme at 65°C for 20 min. Nested PCR reactions was used on the samples to amplify MoTeRs. In the first PCR, MoTeR1 was amplified using ExTaq polymerase (Takara, Madison, WI) following manufactures protocols with the primers MoTeR1001F and MoTeR1001R, and MoTeR2 was amplified using the primers MoTeR2002F and MoTeR2002R. The parameters used in PCR cycling were: 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 60°C for 60 s, and 72°C for 1 min, with a final extension phase of 72°C for 5 min. In the second PCR reaction, 1 µl of the first PCR reaction samples were used with nested PCR primers. For MoTeR1 the primers nMoTeR1001F (5'ATTTTTGCTGACGCCGCA3') and nMoTeR1001R (5'GGCCAATAGGGTGTTTTTCC3') were used in the PCR reactions following protocols outlined above. MoTeR2 was further amplified using the primers nMoTeR2002F (5'GTCCCGCCAAAAATCAAATA3') and nMoTeR2002R (5'TGCTGGACTTTTGGCGTAATAA3'). The following parameters were used in the PCR cycling: 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 56°C for 60 s, and 72°C for 1 min, with a final extension phase of 72°C for 5 min. The PCR amplicons were fractionated on 0.7% agarose gel in 0.5X TBE at 80V for 80 min. After electrophoresis, the agarose gel was stained for 30 min in EtBr and 0.5X TBE. Staining solution was drained, fresh 0.5X TBE was added, and the gel was destained for 20 min.

B.3 Results and Discussion

If MoTeRs were active transposable elements they should express their RNA transcript. Some non-LTRs have poly(A) tails added post-transcriptionally.

However, it was not believed that MoTeR transcript would be polyadenylated. To determine whether MoTeRs were expressed in LpKY97-1A nested reverse transcriptase (RT) PCR reactions were completed. During first strand synthesis either an Oligo(dt)$_{20}$ primer was used or a MoTeR specific primer. This could test whether polyadenylation was occurring if nested PCR of cDNA showed a positive signal in both cDNA created from MoTeR specific primer (MoTeRstartR) and the Oligo(dt)$_{20}$ primer. The results in Figure B-1 show that both MoTeRs were expressed, as bands of expected size were observed in the reactions with RT added. MoTeR1 showed a product only in the reaction where the MoTeR specific primer was used in first strand synthesis, and not with the Oligo(dt)$_{20}$ primer, indicating that this transcript is likely not polyadenylated. MoTeR2 showed PCR products in both the reactions, which could suggest that a poly(A) tract is added post-transcriptionally. However, in MoTeR2 there is a poly(A) tract internally which the Oligo(dt)$_{20}$ could have bound to during reverse transcription. Cloning and sequencing of the PCR amplicons in lanes 2 and 4 in Figure B-1-B confirmed that they were MoTeR2 as expected (data not shown).

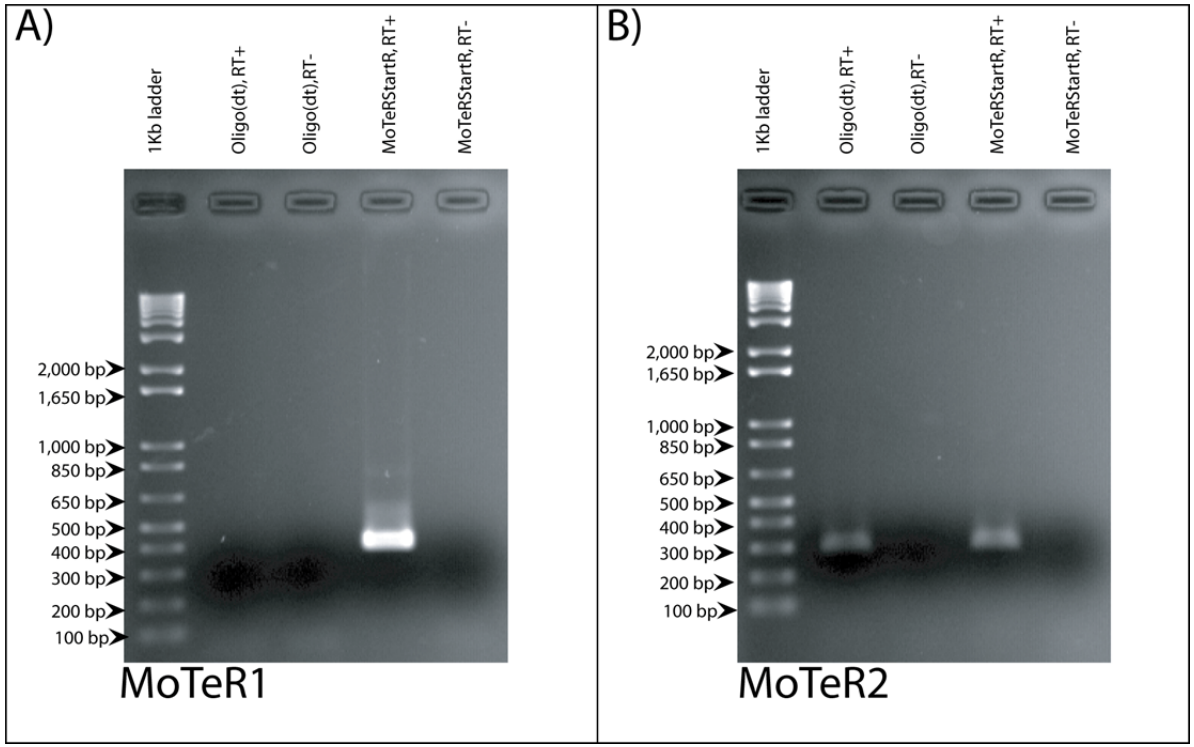

Figure B-1. Nested RT-PCR of MoTeR1 and MoTeR2 of LpKY97-1A RNA. First strand synthesis with reverse transcriptase was used with either an Oligo(dt)$_{20}$ primer or the MoTeRstartR primer from RNA extracted from LpKY971-A, which had been treated with DNase. This was followed by restriction digestion with *Mbo*I to remove any remaining contaminating DNA. Nested PCR was then used to detect expression of MoTeR1 and MoTeR2. The resulting samples were loaded onto a 0.7% agarose gel electrophoresed for 80min at 80V, and stained in ethidium bromide. Results for MoTeR1 are shown in (A), while the results for MoTeR2 are shown in (B) in a nested PCR. Lane one in both (A) and (B) shows 1kB plus DNA ladder. In Lane 2 Oligo(dt)$_{20}$ was used in first stand synthesis and reverse transcriptase was added. In Lane 3 Oligo(dt)$_{20}$ was used in a negative control where no reverse transcriptase was added. In Lane 4 specific was used in first strand synthesis and reverse transcriptase was added. In Lane 5 specific MoTeRstartR was used in a negative control where no reverse transcriptase was added.

Appendix C

LpKY97-1A and IT11 Alignments and Analyses of MoTeR1

Table C-1. Nucleotide Pair Frequencies for LpKY97-1A versus IT11 in MoTeR1 sequence alignments

Domain		ii	si	sv	R	TT	TC	TA	TG	CT	CC	CA	CG	AT	AC	AA	AG	GT	GC	GA	GG	Total
1	Avg	1285	52	27	1.93	311	20	4	3	12	284	4	4	8	3	471	13	1	0	7	219	1364

ii. Identical pairs
si. Transitional pairs
sv. Transversional pairs
R=si/sv

```
#MEGA
!Title LpKY971A and IT11 MoTeR1 Muscle Alignment with variable sites highlighted;
!Format
  DataType=Nucleotide
  NSeqs=2 NSites=1374
 Missing=? Indel=-;

!Domain=Data;
#LpKY97-1A ATCCAAAAAT CGCTTCCATA TACGTATATA GCGGCTTATT TGGACGACGT TTATATTTTA TCCAAAACGC CCGTTAAAGA TAAAATAGCC
#IT11       ATCCAAAAAT CGTTTCCATA TACGTATATA ACGGCTTATT TGGACGACGT TTATATTTTA TCCAAAACGC CCGTTAAAGA TAAAATAGCC

#LpKY97-1A AAAATAATCG AAAAAAGCCC GTTTACCCTA AATTCCGCCA AAACGACAGA AACGGATATC GATACGTTAA AAACCAATGG TTTAAAAACG
#IT11       AAAATAATCG AAAAAAGCCC CTTTACCCTA AATTCCGCCA AAACAACAGA AACGGATATC GAAACCTTAA AAACCAATGG TTTAAAAACG

#LpKY97-1A CTCGGCTCGT TTATTGGACC AACGGAATTA CGGAAGGAAT TTTTGCAAAA TAAAATTCAA AATTTCGAAT CGTCCATTAA CGCCCTGAAA
#IT11       TTAGGTTCGT TTATTGGACC AACGGAATTA CGGAAGGAAT TTTTGCAAAA CAAAATCCAA ACCCTCGAAT TGGCTATAAA CGCCCTTACC

#LpKY97-1A AAACTCCCTA AACAATACGG ATTGCCTAAT CTTGCGTAAA AGTACACAAT TACTTTTACG CCATTTGCTC CGTACTTTAA ATTCCCAGGA
#IT11       AAGCTCCCAA AACAATACGG ATTG -TTAAT TTTACGTAAA AGCACACAAT TATTATTACG CCATTTACTC CGTACATTAA ATTCCCAAGA

#LpKY97-1A CCTGTGGGAA TTATGGGAAA AAACAGATAA ATTAATAGCG GATTTCGTTA TAAATTTAAC TGTTACAAAA CGGAAAAAAC GGCCAATTAC
#IT11       TTTGTGGGAA TTATGGGAAA AAACAGATAA ATTAATAGCG GATTTCGTTA TAAATTTAAC TGTTAAGGA CGGAAAAAAC GCTCAATTAC
```

```
#LpKY97-1A GGATTTCGTT ACGCCGTTAA TTACGTTACC TATAAAGGAC GGAGGTTTTG GATTATTACG GCATAACGGA ATAGCCCAAG ATATTTATTT
#IT11      GGATTTCGTT ACGCCATTAA TTACGTTACC TATAAAGGAC GGAGGTTTTG GAATATTACG CCACAGTGGA ATAGCCCAAG ATATTTATTT

#LpKY97-1A TGCGGCCAAG GATTTAACAA CCGAAATTCG GCACAAAATC CAACGTATAT CCAACGATTT TCCACAAAAT CAAAGCCCTA CCGCCACCGA
#IT11      TGCGGCCATG GATTTAACAG CCGAAATTCG GTATAAAATC CAACGTATAC CCAACGATTT TCCACAAAAT CAAAGCCCTA CCGCCACCGA

#LpKY97-1A GATTTTGCAT TTGTTGCATA ACGGGGTTTT AGCAGATTGC AAAAACGGGT TAACAAACGC CCAATTAAAC GCTTTAACCG AAAACGCTAG
#IT11      AATTTTGCAT TTATTGCATA ACGGGGTTTT AGCAGATTGC AAAAACGGGT TAACAAACGC CCAATTAAAC GCCTTAACCG AAAACGCTAG

#LpKY97-1A TTATTTAGGT CGAAAATGGC TTAACATTTT ACCTATCCAA AAATCAAATC GATTAACGGA TTGGGAAATG GCTGAAGCCG TTCGATTAAG
#IT11      TTATTTAGGT CGAAAATGGC TTAACATTTT ACCTATCCAA AAATCAAATC GATTAACGGA TATAGAAATG GCTGAAGCCG TTCGATCCAG

#LpKY97-1A ATTATTAGCC CCGGTTAAAC CGTTAACCCA CCCCTGCAAC CATTGCGGAA ATCGGACCAA TATAAACCAC GAGGACGTTT GCAAAGGTGC
#IT11      ATTATTAGCC CCGGTTAAAC CGTTAACCCA CCCCTGTAAC CAATGCGGTA ACCGGACCAA TATAAACCAC GAGGACGTTT GCAAAGGTGC

#LpKY97-1A CGTACGCAAA TATACGGCCC GTCACGACCA AATAAACAGA AGTTTCGTCA ATTCGTTAAA AAGTCGACCA GAAATCGACG TCGAAATCGA
#IT11      TGTACGCAAA TATACGGTCC GTCACGACCA AATAAACAAA AGTTTCGTTA ATTCGTTAAA AAGTCGGCCT GAAATCGACG TCGAAATCGA

#LpKY97-1A ACCCGATTTA AATAACGAAA ATAACGT--- - -- ---AAAT AACGCCAATA CAACCACCGA AAATCCCACC CCTAGCCCCA ACGGCCAAAA
#IT11      ACCCGATTTA AATAACGAAA ATAACGTAAA TAACGGAAAT AACGCCAACG CAAACACCGA AAATCCCACC CCTAGCCCCA ACGGCCAAAA

#LpKY97-1A CGATACCGGA TGCCTTTTTA CAACCCCTAT TCGCTCCGGG ACCCGTAACG GCCAAAACGG CCTTAGGGCG GATTTTGCCG TTATTAACGG
#IT11      CGATACCGGA TGCCTTTTTA CAACCCCTAT TCGCTCCGGG ACCCGTAACG GCCAAAACGG CCTTAGGGCG GATTTTGCCG TTATTAACGG

#LpKY97-1A CGTATCCAAA TATTATTACG ACGTGCAAAT CGTTGCAATT AATAAGGATT CCGGTAATAC AAATCCGTTA AATACGTTAG CAGACGCAGC
#IT11      CGTATCCAAA TATTATTACG ACGTGCAAAT CGTTGCAATT AATAAGGATT CCGGTAATAC AAATCCGTTA AATACGTTAA CAGACGCAGC

#LpKY97-1A AAATAACAAA CGACGTAAAT ACCAATTTTT GGATCCATTT TTCCATCCAA TTATAATAAG CGCCGGAGGC CTTATGGAAA AGGATACAGC
#IT11      AAATAACAAA CGACGTAAAT ACCAATTTTT GGACCCTTTT TTCCATCCAA TTATAATAAG CGCCGGAGGC CTTATGGAAA AAAATACAGC

#LpKY97-1A ACAGGCGTAC AAACAAATCC AAAA
#IT11      ACAGGCTTAT AAACAAATTC AAAA
```

References

ABAD, J. P., B. DE PABLOS, K. OSOEGAWA, P. J. DE JONG, A. MARTIN-GALLARDO *et al.*, 2004a Genomic analysis of *Drosophila melanogaster* telomeres: Full-length copies of HeT-A and TART elements at telomeres. Molecular Biology and Evolution **21:** 1613-1619.

ABAD, J. P., B. DE PABLOS, K. OSOEGAWA, P. J. DE JONG, A. MARTIN-GALLARDO *et al.*, 2004b TAHRE, a novel telomeric retrotransposon from *Drosophila melanogaster*, reveals the origin of Drosophila telomeres. Molecular Biology and Evolution **21:** 1620-1624.

AHN, S. W., and S. H. OU, 1982 Quantitative resistance of rice to blast disease. Phytopathology **72:** 279-282.

AKAGI, Y., H. AKAMATSU, H. OTANI and M. KODAMA, 2009 Horizontal chromsome transfer, a mechanism for the evolution and differentation of a plant-pathogenic fungus. Eukaryotic Cell **8:** 1732-1738.

AKSOY, S., S. WILLIAMS, S. CHANG and F. F. RICHARDS, 1990 SLACS retrotransposon from *Trypanosoma brucei gambiense* is similar to mammalian LINEs. Nucleic Acids Research **18:** 785-792.

ALIZON, M., S. WAINHOBSON, L. MONTAGNIER and P. SONIGO, 1986 Genetic-variability of the AIDs virus - nucleotide-sequence analysis of 2 isolates from african patients. Cell **46:** 63-74.

AMARGER, V., D. GAUGUIER, M. YERLE, F. APIOU, P. PINTON *et al.*, 1998 Analysis of distribution in the human, pig, and rat genomes points toward a general subtelomeric origin of minisatellite structures. Genomics **52:** 62-71.

AMASINO, R. M., 1986 Acceleration of nucleic acid hybridization rate by polyethylene glycol. Analytical Biochemistry **152:** 304-307.

ANAYA, N., and M. I. RONCERO, 1995 Skippy, a retrotransposon from the fungal plant pathogen *Fusarium oxysporum*. Molecular and General Genetics **249:** 637-647.

ANZAI, T., M. OSANAI, M. HAMADA and H. FUJIWARA, 2005 Functional roles of 3′-terminal structures of template RNA during in vivo retrotransposition of non-LTR retrotransposon, R1Bm. Nucleic Acids Research **33:** 1993-2002.

ANZAI, T., H. TAKAHASHI and H. FUJIWARA, 2001 Sequence-specific recognition and cleavage of telomeric repeat (TTAGG)n by endonuclease of non-Long Terminal Repeat retrotransposon TRAS1. Molecular and Cellular Biology **21:** 100-108.

ARKHIPOVA, I. R., 2005 Mobile genetic elements and sexual reproduction. Cytogenetic Genome Research **110:** 372-382.

ARKHIPOVA, I. R., and M. MESELSON, 2000 Trasposable elements in sexual and ancient asexual taxa. Proceedings of the National Academy of Sciences of the United States of America **97:** 14473-14477.

ARKHIPOVA, I. R., K. I. PYATKOV, M. MESELSON and M. B. EVGEN'EV, 2003 Retroelements containing introns in diverse invertebrate taxa. Nature Genetics **33:** 123-124.

ASHIZAWA, T., K. ZENBAYASHI and S. KOIZUMI, 1999 Effect of partial restance on rice blast control with cultivar mixtures. Phytopathology **89:** S3.

ATILANO, R. A., and P. BUSEY, 1983 Susceptibility of St. Augustinegrass germplasm to *Pyricularia grisea*. Plant Disease **67:** 782-783.

BAIN, D. C., B. M. PATEL and M. V. PATEL, 1972 Blast of ryegrass in Mississippi. Plant Disease Report **56:** 210.

BARRY, J. D., M. L. GINGER, P. BURTON and R. MCCULLOCH, 2003 Why are parasite contingency genes often associated with telomeres? International Journal for Parasitology **33:** 29-45.

BASSHAM, S., A. BEAM and J. SHAMPAY, 1998 Telomere variation in *Xenopus laevis*. Molecular and Cellular Biology **18:** 269-275.

BAUR, J. A., Y. ZOU, J. W. SHAY and W. E. WRIGHT, 2001 Telomere position effect in human cells. Science **292:** 2075-2077.

BENDER, W., M. , F. K. AKAM, P. BEACHY, M. PEIFER, P. SPIERER *et al.*, 1983 Molecular genetics of the bithorax complex of Drosophila melanogaster. Science **221:** 23-29.

BENSON, G., 1999 Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Research **27:** 573-580.

BERRIMAN, M., E. GHEDIN, C. HERTZ-FOWLER, G. BLANDIN, H. RENAULD *et al.*, 2005 The genome of the African trypanosome *Trypanosoma brucei*. Science **309:** 416-422.

BESANSKY, N. J., J. A. BEDELL and O. MUKABAYIRE, 1994 Q: a new retrotransposon from the mosquito *Anopheles gambiae*. Insect Molecular Biology **3:** 49-56.

BESANSKY, N. J., S. M. PASKEWITZ, D. M. HAMM and F. H. COLLINS, 1992 Distinct families of site-specific retrotransposons occupy identical positions in the rRNA genes of *Anopheles gambiae*. Molecular and Cellular Biology **12:** 5102-5110.

BIBILLO, A., and T. H. EICKBUSH, 2004 End-to-end template jumping by the reverse transcriptase encoded by the R2 retrotransposon. Journal of Biological Chemistry **279:** 14945–14953.

BIESSMANN, H., and J. M. MASON, 1997 Telomere maintenance without telomerase. Chromosoma **106:** 63-69.

BIESSMANN, H., K. VALGEIRSDOTTIR, A. LOFSKY, C. CHIN, B. GINTHER *et al.*, 1992 HeT-A, a transposable element specifically involved in "healing" broken chromosome ends in *Drosophila melanogaster*. Molecular and Cellular Biology **12:** 3910-3918.

BLACKBURN, E. H., and K. COLLINS, 2011 Telomerase: An RNP Enzyme Synthesizes DNA. Cold Spring Harbor Perspectives in Biology **3:** 1-9.

BLASCO, M. A., H. W. LEE, M. P. HANDE, E. SAMPER, P. M. LANSDORP *et al.*, 1997 Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. Cell **91:** 25-34.

BLESA, D., and M. J. MARTINEZ-SEBASTIAN, 1997 Bilbo, a non-LTR retrotransposon of *Drosophila subobscura*: a clue to the evolution of LINE-like elements in Drosophila. Molecular Biology and Evolution **14:** 1145-1153.

BLINOV, A. G., Y. V. SOBANOV, S. S. BOGACHEV, A. P. DONCHENKO and M. A. FILIPPOVA, 1993 The *Chironomus thummi* genome contains a non-LTR retrotransposon. Molecular and General Genetics **237:** 412-420.

BODEM, J., M. LOCHELT, H. DELIUS and R. M. FLUGEL, 1998 Detection of subgenomic cDNAs and mapping of feline foamy virus mRNAs reveals complex patterns of transcription. Virology **244:** 417-426.

BOLTON, M. D., and B. P. H. J. TOMMA, 2008 The complexity of nitrogen metabolism and nitrogen-regulated gene expression in plant pathogenic fungi. Physiology of Molecular Plant Pathology **72:** 104-110.

BONMAN, J. M., 1992 Durable resistance to rice blast disease-environmental influences. Euphytica **63:** 115-123.

BONMAN, J. M., B. A. ESTRADA, C. K. KIM, D. S. RA and E. J. LEE, 1991 Assessment of blast disease and yield loss in susceptible and partially resistant rice cultivars in two irrigated lowland environments. Plant Disease **75:** 462-466.

BONOS, S., C. KUBIK, B. B. CLARKE and W. A. MEYER, 2004 Breeding perennial regrass for resistance to gray leaf spot. Crop Science **44:** 575-580.

BONOS, S., D. RUSH, K. HIGNIGHT, S. LANGLOIS and W. A. MEYER, 2005 The effect of selection on gray leaf spot resistance in perennial ryegrass. International Turfgrass Society Research Journal **10:** 501-507.

BORROMEO, E. S., R. J. NELSON, J. M. BONMAN and H. LEUNG, 1993 Genetic differentiation among isolates of Pyricularia infecting rice and weed hosts. Phytopathology **83:** 393-399.

BRADSHAW, H., 1996 (Accession no. T25782, GI 7503675). GenBank Direct Submission.

BRAUMANN, I., M. VAN DEN BERG and F. KEMPKEN, 2008 Repeat induced point mutation in two asexual fungi, *Aspergillus niger* and *Penicillium chrysogenum*. Current Genetics **53:** 287-297.

BROUHA, B., J. SCHUSTAK, R. M. BADGE, S. LUTZ-PRIGGE, A. H. FARLEY *et al.*, 2003 Hot L1s account for the bulk of retrotransposition in the human population. Proceedings of the National Academy of Sciences of the United States of America **100:** 5280-5285.

BROUN, P., M. W. GANAL and S. D. TANKSLEY, 1992 Telomeric arrays display high levels of heritable polymorphism among closely related plant varieties. Proceedings of the National Academy of Sciences of the United States of America **89:** 1354-1357.

BURCH, J. B., D. L. DAVIS and N. B. HAAS, 1993 Chicken repeat 1 elements contain a pol-like open reading frame and belong to the non-long terminal repeat class of retrotransposons. Proceedings of the National Academy of Sciences of the United States of America **90:** 8199-8203.

BURKE, W. D., C. C. CALALANG and T. H. EICKBUSH, 1987 The site-specific ribosomal insertion element type II of *Bombyx mori* (R2Bm) contains the coding sequence for a reverse transcriptase-like enzyme. Molecular and Cellular Biology **7:** 2221-2230.

BURKE, W. D., H. S. MALIK, J. P. JONES and T. H. EICKBUSH, 1999 The domain structure and retrotransposition mechanism of R2 elements are conserved throughout arthropods. Molecular Biology and Evolution **16:** 502-511.

BURKE, W. D., H. S. MALIK, S. M. RICH and T. H. EICKBUSH, 2002 Ancient lineages of non-LTR retrotransposons in the primitive eukaryote, *Giardia lamblia*. Molecular Biology and Evolution **19:** 619-630.

BURKE, W. D., F. MULLER and T. H. EICKBUSH, 1995 R4, a non-LTR retrotransposon specific to the large subunit rRNA genes of nematodes. Nucleic Acids Research **23:** 4628-4634.

CAMBARERI, E. B., J. HELBER and J. A. KINSEY, 1994 Tad1-1, an active LINE-like element of *Neurospora crassa*. Molecular and General Genetics **242:** 658-665.

CAMBARERI, E. B., B. C. JENSEN, E. SCHABTACH and E. U. SELKER, 1989 Repeat-induced G-C to A-T mutations in Neurospora. Science **244:** 1571-1575.

CARLSON, M., J. L. CELENZA and F. J. ENG, 1985 Evolution of the dispersed SUC gene family of Saccharomyces by rearrangements of chromosome telomeres. Molecular Cell Biology **5:** 2894-2902.

CASACUBERTA, E., and M.-L. PARDUE, 2005 HeT-A and TART, two drosophila retrotransposons with a bona fide role in chromosome structure for more than 60 million years. Cytogenetic Genome Research **110:** 152-159.

CASTANO, J., D. R. MACKENZIE and R. R. NELSON, 1989 Components analysis of race non-specific resistance to blast disease of rice caused by *Pyricularia oryzae*. Journal of Phytopathology**:** 89-99.

CHABOISSIER, M. C., D. J. FINNEGAN and A. BUCHETON, 2000 Retrotransposition of the I factor, a non-long terminal repeat retrotrasposon of Drosophila, generates tandem repeats at the 3' end. Nucleic Acids Research **28:** 2467-2472.

CHADWICK, B. J., R. J. COELEN, G. E. WILCOX, L. M. SAMMELS and G. KERTAYADNYA, 1995 Nucleotide-sequence analysis of jembrana disease virus - a bovine lentivirus associated with an acute disease syndrome. Journal of General Virology **76:** 1637-1650.

CHAPMAN, K., A. BYSTROM and B. J., 1992 Initiator methionine tRNA is essential for Ty1 transposition in yeast. Proceedings of the National Academy of Sciences of the United States of America **89:** 3236-3240.

CHARRON, M. J., E. READ, S. R. HAUT and C. A. MICHELS, 1989 Molecular evolution of the telomere-associated *Mal* loci of Saccharomyces. Genetics **122:** 307-316.

CHAUHAN, R. S., M. L. FARMAN, H. B. ZHANG and S. A. LEONG, 2002 Genetic and physical mapping of a rice blast resistance locus, Pi-CO39(t). Molecular Genetics and Genomics **267:** 603-612.

CHEN, H., S. P. WANG, Y. Z. XING, C. G. XU, P. M. HAYES *et al.*, 2003 Comparative analyses of genomic locations and race specificities of loci for quantitative resistance to *Pyricularia grisea* in rice and barley. Proceedings of the National Academy of Sciences of the United States of America **100:** 2544-2549.

CHOI, Y., N. ISHIGURO, M. SHINAGAWA, C. J. KIM, Y. OKAMOTO *et al.*, 1999 Molecular structure of canine LINE-1 elements in canine transmissible venereal tumor. Animal Genetics **30:** 51-53.

CHRISTENSEN, S. M., A. BIBILLO and T. H. EICKBUSH, 2005 Role of the *Bombyx mori* R2 element N-terminaldomain in the target-primed reverse transcription(TPRT) reaction. Nucleic Acids Research **33:** 6461-6468.

CHRISTENSEN, S. M., J. YE and T. H. EICKBUSH, 2006 RNA from the 5' end of the R2 retrotransposon controls R2 protein binding to and cleavage of its DNA target site. Proceedings of the National Academy of Sciences of the United States of America **103:** 17602-17607.

CHUMA, I., Y. HOTTA and Y. TOSA, 2011a Instability of subtelomeric regions during meiosis in *Magnaporthe oryzae*. Journal of General Plant Pathology **77:** 317-325.

CHUMA, I., C. ISOBE, Y. HOTTA, K. IBARAGI, N. FUTAMATA *et al.*, 2011b Multiple translocation of the AVR-Pita effector gene among chromosomes of the rice blast fungus *Magnaporthe oryzae* and related species. PLoS Pathogens **7:** e1002147.

CHUMA, I., T. SHINOGI, N. HOSOGI, K.-I. IKEDA, H. NAKAYASHIKI *et al.*, 2009 Cytological characteristics of microconidia of *Magnaporthe oryzae*. Journal of General Plant Pathology **75:** 353-358.

CHUMLEY, F. G., and B. VALENT, 1990 Genetic analysis of the Melanin-deficient nonpathogenic mutant of *Magnaporthe grisea*. Molecular Plant-Microbe Interactions **3:** 135-145.

COHEN, J. I., R. H. MILLER, B. ROSENBLUM, K. DENNISTON, J. L. GERIN *et al.*, 1988 Sequence comparison of woodchuck hepatitis virus replicative forms shows conservation of the genome. Virology **162:** 12-20.

COLEMAN, J. J., S. D. ROUNSLEY, M. RODRIGUEZ-CARRES, A. KUO, C. C. WASMANN *et al.*, 2009 The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. PLoS Genetics **5:** e1000618.

COLLINS, M., and G. M. RUBIN, 1982 Structure of the Drosophila mutable allele, white-crimson, and its white-ivory and wild-type derivatives. . Cell **30:** 71-79.

COPPIN, E., R. DEBUCHY, S. ARNAISE and M. PICARD, 1997 Mating types and sexual development in filamentous ascomycetes. Microbiology and Molecular Biology Reviews **61:** 411-428.

CORCORAN, L. M., J. K. THOMPSON, D. WALLIKER and D. J. KEMP, 1988 Homologous recombination within subtelomeric repeat sequences generates chromosome size polymorphisms in *P. falciparum*. Cell **53:** 807-813.

CORREA-VICTORIA, F. J., R. S. ZEIGLER and M. LEVY, 1994 Virulence characteristics of genetic families of *Pyricularia grisea* in Columbia, pp. 211-229 in *Rice Disease*,

edited by R. S. ZEIGLER, S. A. LEONG and P. S. TENG. CAB International, Wallingford, UK.

COUCH, B. C., I. FUDAL, M.-H. LEBRUM, D. THARREAU, B. VALENT *et al.*, 2005 Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. Genetics **170:** 613-630.

COUCH, B. C., and L. M. KOHN, 2002 A multilocus gene genealogy concordant with host preference indicates segregation of a new species, *Magnaporthe oryzae*, from *M. grisea*. Mycologia **94:** 683-693.

COULSTON, J., H. NAIF, R. BRANDON, S. KUMAR, S. KHAN *et al.*, 1990 Molecular-cloning and sequencing of an australian isolate of proviral bovine leukemia-virus DNA - comparison with other isolates. Journal of General Virology **71:** 1737-1746.

COVERT, S. F., 1998 Supernumerary chromosome in filamentous fungi. Current Genetics **33:** 311-319.

CUMMINGS, D. J., I. A. MACNEIL, J. DOMENICO and E. T. MATSUURA, 1985 Excision-amplification of mitochondrial DNA during senescence in *Podospora anserina*. DNA sequence analysis of three unique "plasmids". Journal of Molecular Biology **185:** 659-680.

CURCIO, M. J., and K. M. DERBYSHIRE, 2003 The outs and ins of transposition: from mu to kangaroo. Nature Reviews Molecular Cell Biology **4:** 865-877.

DABOUSSI, M. J., 1997 Fungal transposable elements and genome evolution. Genetica **100:** 253-260.

DAWSON, A., E. HARTSWOOD, T. PATERSON and D. J. FINNEGAN, 1997 A LINE-like transposable element in Drosophila, the I factor, encodes a protein with properties similar to those of retroviral nucleocapsids. The EMBO Journal **16:** 4448-4455.

DE KOCHKO, A., B. VERDAGUER, N. TAYLOR, R. CARCAMO, R. N. BEACHY *et al.*, 1998 Cassava vein mosaic virus (CsVMV), type species for a new genus of plant double stranded DNA viruses? Archives of Virology **143:** 945-962.

DE LANGE, T., 2005 Shelterin: the protein complex that shapes and safeguards human telomeres. Genes & Development **19:** 2100-2110.

DEAN, R. A., N. J. TALBOT, D. J. EBBOLE, M. L. FARMAN, T. K. MITCHELL *et al.*, 2005 The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature **434:** 980-986.

DERNOEDEN, P. H., 1996 Perennial ryegrass and gray leaf spot. Golf Course Management **64:** 49-52.

DI NOCERA, P. P., and G. CASARI, 1987 Related polypeptides are encoded by Drosophila F elements, I factors, and mammalian L1 sequences. Proceedings of the National Academy of Sciences of the United States of America **84:** 5843-5847.

DIOH, W., D. THARREAU, J. L. NOTTEGHEM, M. ORBACH and M.-H. LEBRUN, 2000 Mapping of avirulence genes in the rice blast fungus, *Magnaporthe grisea*, with RFLP and RAPD Markers. Molecular Plant-Microbe Interactions **13:** 217-227.

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 Interactions **6:** 114-126.

DOMBROSKI, B. A., S. L. MATHIAS, E. NANMAKMAR, A. F. SCOTT and H. H. KAZAZIAN, 1991 Isolation of an active human transposable element. Science **254:** 1805-1808.

DORING, P., and P. STARLINGER, 1984 Barbara McClintock's controlling elements: now at the DNA level. Cell **39:** 253-259.

DOUHAN, G. W., K. A. DE LA CERDA, C. A. GREER and F. P. WONG, 2011 Contrasting genetic structure between *Magnaporthe grisea* populaitons associated with golf course turfgrasses *Lolium perenne* (Perennial Ryegrass) and *Pennisetum clandestinum* (Kikuyugrass). Phytopathology **101:** 85-91.

DREW, A. C., and P. J. BRINDLEY, 1997 A retrotransposon of the non-long terminal repeat class from the human blood fluke *Schistosoma mansoni*. Similarities to the chicken-repeat-1-like elements of vertebrates. Molecular Biology and Evolution **14:** 602-610.

DUFRESNE, M., and A. E. OSBOURN, 2001 Definition of tissue-specific and general requirements for plant infection in a phytopathogenic fungus. Molecular Plant-Microbe Interactions **14:** 300-307.

DUVERNELL, D. D., and B. J. TURNER, 1998 Swimmer 1, a new low-copy-number LINE family in teleost genomes with sequence similarity to mammalian L1. Molecular Biology and Evolution **15:** 1791-1793.

EICHLER, E. E., and D. SANKOFF, 2003 Sturctural dynamics of eukaryototic chromosome evolution. Science **301:** 793-797.

EICKBUSH, D. G., and T. H. EICKBUSH, 1995 Vertical transmission of the retrotransposable elements R1 and R2 during the evolution of the *Drosophila melanogaster* species subgroup. Genetics **139:** 671-684.

EICKBUSH, D. G., and T. H. EICKBUSH, 2012 R2 and R2/R1 hybrid non-autonomous retrotransposons derived by internal deletions of full-length elements. Mobile DNA **3:** 1-15.

EICKBUSH, T. H., 2002 R2 and related site-specific non-LTR retrotransposons, pp. 813-835 in *Mobile DNA II*, edited by N. CRAIG, R. CRAIGIE, M. GELLERT and A. LAMBOWITZ. American Society of Microbiology Press, Washington D.C.

EICKBUSH, T. H., and H. S. MALIK, 2002 Evolution of retrotransposons, pp. 1111-1144 in *Mobile DNA II*, edited by N. CRAIG, R. CRAIGIE, M. GELLERT and A. LAMBOWITZ. American Society of Microbiology Press, Washington, D.C.

ENGELS, W. R., D. M. JOHNSON-SCHLITZ, W. B. EGGLESTON and J. SVED, 1990 High-frequency P element loss in Drosophila is homolog dependent. Cell **62:** 515-525.

ETO, Y., K. IKEDA, I. CHUMA, T. KATAOKA, S. KURODA *et al.*, 2001 Comparative analyses of the distribution of various transposable elements in Pyricularia and their activity during and after the sexual cycle. Molecular and General Genetics **264:** 565-577.

EVGENEV, M. B., H. ZELENTSOVA, N. SHOSTAK, M. KOZITSINA, V. BARSKYI *et al.*, 1997 Penelope, a new family of transposable elements and its possible role in hybrid dysgenesis in *Drosophila virilis*. Proceedings of the National Academy of Sciences of the United States of America **94:** 196-201.

FARMAN, M. L., 2002 *Pyricularia grisea* isolates causing gray leaf spot on perennial ryegrass (*Lolium perenne*) in the United States: relationship to *P. grisea* isolates from other host plants. (Erratum: Nov 2002, v. 92 (11), p. 1253.). Phytopathology **92:** 245-254.

FARMAN, M. L., 2007 Telomeres in the rice blast fungus *Magnaporthe oryzae*: the world of the end as we know it. FEMS Microbiology Letters **273:** 125-132.

FARMAN, M. L., 2011 Targeted cloning of fungal telomeres. Methods in molecular biology **722:** 11-31.

FARMAN, M. L., Y. ETO, T. NAKAO, Y. TOSA, H. NAKAYASHIKI *et al.*, 2002 Analysis of the structure of the AVR1-CO39 avirulence locus in virulent rice-infecting isolates of *Magnaporthe grisea*. Molecular Plant-Microbe Interactions **15:** 6-16.

FARMAN, M. L., and Y. S. KIM, 2005 Telomere hypervariability in *Magnaporthe oryzae*. Molecular Plant Pathology **6:** 287-298.

FARMAN, M. L., and S. A. LEONG, 1995 Genetic and physical mapping of telomeres in the rice blast fungus, *Magnaporthe grisea*. Genetics **140:** 479-492.

FARMAN, M. L., and S. A. LEONG, 1998 Chromosome walking to the AVR1-CO39 avirulence gene of Magnaporthe grisea: Discrepancy between the physical and genetic maps. Genetics **150:** 1049-1058.

FARMAN, M. L., S. TAURA and S. A. LEONG, 1996a The *Magnaporthe grisea* DNA fingerprinting probe MGR586 contains the 3' end of an inverted repeat transposon. Molecular and General Genetics **251:** 675-681.

FARMAN, M. L., Y. TOSA, N. NITTA and S. A. LEONG, 1996b MAGGY, a retrotransposon in the genome of the rice blast fungus *Magnaporthe grisea*. Molecular and General Genetics **251:** 665-674.

FAWCETT, D. H., C. K. LISTER, E. KELLET and D. J. FINNEGAN, 1986 Transposable elements controlling I-R hybrid dysgenesis in *D. melanogaster* are similar to mammalian LINEs. Cell **47:** 1007-1015.

FELLERHOFF, B., F. ECKARDT-SCHUPP and A. A. FRIEDL, 2000 Subtelomeric repeat amplification is assocated with growth at elvated temperature in yku70 mutants of *Sacchromyces cerevisiae*. Genetics **154:** 1039-1051.

FELSENSTEIN, J., 1985 Confidence-limits on phylogenies - an approach using the bootstrap. Evolution **39:** 783-791.

FENG, Q., J. V. MORAN, H. H. KAZAZIAN, JR. and J. D. BOEKE, 1996 Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell **87:** 905-915.

FENG, Y., S. MOORE, D. GARFINKEL and A. REIN, 2000 The genomic RNA in Ty1 virus-like particles is dimeric. Journal of Virology **74:** 10819-10821.

FERREIRA, M. G., K. M. MILLER and J. P. COOPER, 2004 Indecent exposure: When telomeres become uncapped. Molecular Cell **13:** 7-18.

FINNEGAN, D. J., 1989 Eukaryotic Transposable Elements and Genome Evolution. Trends in Genetics **5:** 103-107.

FITZGERALD, M. S., K. RIHA, F. GAO, S. X. REN, T. D. MCKNIGHT *et al.*, 1999 Disruption of the telomerase catalytic subunit gene from Arabidopsis inactivates telomerase and leads to a slow loss of telomeric DNA. Proceedings of the National Academy of Sciences of the United States of America **96:** 14813-14818.

FLOR, H. H., 1971 Current status of gene-for-gene concept. Annual Review of Phytopathology **9:** 275-296.

FOURCADE-PERONNET, F., L. D'AURIOL, J. BECKER, F. GALIBERT and M. BEST-BELPOMME, 1988 Primary structure and functional organization of Drosophila 1731 retrotransposon. Nucleic Acids Research **16:** 6113-6125.

FRASER, M. L., 1996 Susceptibility of tall fescues to gray leaf spot. Biological and Cultural Tests **12:** 130.

FREEMAN, T. E., 1962 Gray leaf spot preys on St. Augustinegrass. Florida Agriculture Experiment Station State Research Report **7:** 8-9.

FUDAL, I., H. U. BOHNERT, D. THARREAU and M. H. LEBRUN, 2005 Transposition of MINE, a composite retrotransposon, in the avirulence gene ACE1 of the rice blast fungus *Magnaporthe grisea*. Fungal Genetics and Biology **42:** 761-772.

FUDAL, I., S. ROSS, H. BRUN, A. L. BESNARD, M. ERMEL *et al.*, 2009 Repeat-induced point mutation (RIP) as an alternative mechanism of evolution toward virulence in *Leptosphaeria maculans*. Molecular Plant-Microbe Interactions **22:** 932-941.

FUJIWARA, H., M. OSANAI, T. MATSUMOTO and K. K. KOJIMA, 2005 Telomere-specific non-LTR retrotransposons and telomere maintenance in the silkworm, *Bombyx mori*. Chromosome Research **13:** 455-467.

GABRIEL, A., T. J. YEN, D. C. SCHWARTZ, C. L. SMITH, J. D. BOEKE *et al.*, 1990 A rapidly rearranging retrotransposon within the miniexon gene locus of *Crithidia fasciculata*. Molecular and Cellular Biology **10:** 615-624.

GARRETT, J. E., D. S. KNUTZON and D. CARROLL, 1989 Composite transposable elements in the *Xenopus laevis* genome. Molecular and Cellular Biology **9:** 3018-3027.

GARVEY, K. J., M. S. OBERSTE, J. E. ELSER, M. J. BRAUN and M. A. GONDA, 1990 Nucleotide-sequence and genome organization of biologically-active proviruses of the bovine immunodeficiency-like virus. Virology **175:** 391-409.

GILBERT, N., and D. LABUDA, 1999 CORE-SINEs: eukaryotic short interspersed retroposing elements with common sequence motifs. Proceedings of the National Academy of Sciences of the United States of America **96:** 2869-2874.

GLADYSHEV, E. A., and I. R. ARKHIPOVA, 2007 Telomere-associated endonuclease-deficient Penelope-like retroelements in diverse eukaryotes. Proceedings of the National Academy of Sciences of the United States of America **104:** 9352-9357.

GLADYSHEV, E. A., and I. R. ARKIPOVA, 2010 A subtelomeric non-LTR retrotransposon *Hebe* in the bdelloid rotifer *Adineta vaga* is subject of inactivation by deletions but not 5' truncations. Mobile DNA **1:** 1-12.

GOLUBOVSKY, M., A. KONEV, M. WALTER, H. BIESSMANN and J. MASON, 2001 Terminal retrotransposons activate a subtelomeric white transgene at the 2L telomere in Drosophila. Genetics **158:** 1111-1123.

GOODWIN, T. J., and R. T. POULTER, 2001 The diversity of retrotransposons in the yeast *Cryptococcus neoformans*. Yeast **18:** 865-880.

GOODWIN, T. J. D., and R. T. M. POULTER, 2005 Penelope-like retroelements in fungi (Accession no. AAX11377.1). GenBank Direct Submission.

GOTTSCHLING, D. E., O. M. APARICIO, B. L. BILLINGTON and V. A. ZAKAIN, 1990 Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. Cell **64:** 751-762.

GOUT, L., M. L. KUHN, L. VINCENOT, S. BERNARD-SAMAIN, L. CATTOLICO *et al.*, 2007 Genome structure impacts molecular evolution at the AvrLm1 avirulence locus of the plant pathogen *Leptosphaeria maculans*. Environmental Microbiology **9:** 2978-2992.

GRANDBASTIEN, M. A., A. SPIELMANN and M. CABOCHE, 1989 Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature **337:** 376-380.

GRANT, D., and R. SHOEMAKER, 1997 Molecular hybridization, pp. 15-26 in *DNA markers: Protocols, Applications and Overviews*, edited by G. CAETANO-ANOLLES and P. M. GRESSHOFF. Wiley-VCH, New York.

GREIDER, C. W., and E. H. BLACKBURN, 1987 The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. Cell **51:** 887-898.

GRIFFITH, J. D., L. COMEAU, S. ROSENFIELD, R. M. STANSEL and A. BIANCHI, 1999 Mammalian telomeres end in a large duplex loop. Cell **97:** 503-514.

HAAS, N. B., J. M. GRABOWSKI, J. NORTH, J. V. MORAN, H. H. KAZAZIAN *et al.*, 2001 Subfamilies of the CR1 non-LTR retrotransposons have different 5'UTR sequences but are otherwise conserved. Gene **265:** 175-183.

HAGEN, L. S., M. JACQUEMOND, A. LEPINGLE, H. LOT and M. TEPFER, 1993 Nucleotide sequence and genomic organization of cacao swollen shoot virus. Virology **196:** 619-628.

HAMER, J. E., L. FARRALL, M. J. ORBACH, B. VALENT and F. G. CHUMLEY, 1989a Host species-specific conservation of a family of repeated DNA sequences in the genome of a fungal plant pathogen. Proceedings of the National Academy of Sciences of the United States of America **86:** 9981-9985.

HAMER, J. E., B. VALENT and F. G. CHUMLEY, 1989b Mutations at SMO genetic locus affect the shape of diverse cell types in the rice blast fungus. Genetics **122:** 351-361.

HAN, J. S., 2010 Non-long terminal repeat (non-LTR) retrotransposons: mechanisms, recent developments, and unanswered questions. Mobile DNA **1:** 1-12.

HAN, Y., X. LIU, U. BENNY, C. KISTLER and H. VANETTEN, 2001 Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca*. Plant Journal **25:** 305-314.

HANSEN, L. J., D. L. CHALKER and S. B. SANDMEYER, 1988 Ty3, a yeast retrotransposon associated with tRNA genes, has homology to animal retroviruses. Molecular and Cellular Biology **8:** 5245-5256.

HARPER, G., and R. HULL, 1998 Cloning and sequence analysis of banana streak virus DNA. Virus Genes **17:** 271-278.

HARTL, D. L., A. R. LOHE and E. R. LOZOVSKAYA, 1997 Modern thoughts on an ancyent marinere: function, evolution, regulation. Annual Review of Genetics **31:** 337-358.

HATTA, R., K. ITO, Y. HOSAKI, T. TANAKA, A. TANAKA *et al.*, 2002 A conditionally dispensable chromosome controls host-specific pathogenicity in the fungal plant pathogen *Alternaria alternata*. Genetics **161:** 59-70.

HAVECKER, E. R., X. GAO and D. F. VOYTAS, 2004 The diversity of LTR retrotransposons. Genome Biology **5:** 225.

HE, C., A. RUSU, A. POPLAWSKI, J. IRWIN and J. MANNERS, 1998 Transfer of a supernumerary chromosome between vegetatively incompatible biotypes of the fungus *Colletotrichum gloeosporioides*. Genetics **150:** 1459-1466.

HERR, W., 1984 Nucleotide sequence of AKV murine leukemia virus. Journal of Virology **49:** 471-478.

HERZER, P. J., S. INOUYE and M. INOUYE, 1992 Retron-Ec107 is inserted into the *Escherichia coli* genome by replacing a palindromic 34bp intergenic sequence. Molecular Microbiology **6:** 345-354.

HIGASHIYAMA, T., Y. NOUTOSHI, M. FUJIE and T. YAMADA, 1997 Zepp, a LINE-like retrotransposon accumulated in the Chlorella telomeric region. The EMBO Journal **16:** 3715-3723.

HOFMANN, N., and A. HABLIN, 2000 Progress in identifying resistance to gray leaf spot in perennial ryegrass., pp. 161-162 in *Annual Meeting of the American Society of Agronomy*. American Society of Agronomy, Minneapolis, MN.

HOROWITZ, H., and J. E. HABER, 1985 Identification of autonomously replicating circular subtelomeric Y' elements in *Saccharomyces cerevisiae*. Molecular and Cellular Biology **5:** 2369-2380.

HOROWITZ, H., P. THORBURN and J. E. HABER, 1984 Rearrangements of highly polymorphic regions near telomeres of *Saccharomyces cerevisiae*. Molecular Cell Biology **4:** 2509-2517.

HOVEL-MINER, G. A., C. E. BOOTHROYD, M. MUGNIER, O. DREESEN, G. A. M. CROSS *et al.*, 2012 Telomere length affects the frequency and mechanism of antigenic variation in *Trypanosoma brucei*. PLoS Pathogens **8:** e1002900.

HOWARD, R. J., 1994 Cell biology of pathogenesis, pp. 3-22 in *Rice Blast Disease*, edited by R. S. ZEIGLER, S. A. LEONG and P. S. TENG. CAB International, Oxon, UK.

HOWARD, R. J., M. A. FERRARI, D. H. ROACH and N. P. MONEY, 1991 Penetration of hard substrates by a fungus employing enormous turgor pressures. Proceedings of the National Academy of Sciences of the United States of America **88:** 11281-11284.

IKEDA, K., 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.

IKEDA, K., H. NAKAYASHIKI, M. TAKAGI, Y. TOSA and S. MAYAMA, 2001 Heat shock, copper sulfate and oxidative stress activate the retrotransposon MAGGY resident in the plant pathogenic fungus *Magnaporthe grisea*. Molecular Genetics and Genomics **266:** 318-325.

INOUYE, S., P. J. HERZER and M. INOUYE, 1990 Two independent retrons with highly diverse reverse transcriptases in *Myxococcus xanthus*. Proceedings of the National Academy of Sciences of the United States of America **87:** 942-945.

INOUYE, S., M. Y. HSU, S. EAGLE and M. INOUYE, 1989 Reverse transcriptase associated with the biosynthesis of the branched RNA-linked msDNA in *Myxococcus xanthus*. Cell **56:** 709-717.

INOUYE, S., S. YUKI and K. SAIGO, 1986 Complete nucleotide sequence and genome organization of a Drosophila transposable genetic element, 297. European Journal of Biochemistry **154:** 417-425.

IRELAN, J. T., A. T. HAGEMANN and E. U. SELKER, 1994 High frequency repeat-induced point mutation (RIP) is not associated with efficient recombination in Neurospora. Genetics **138:** 1093-1103.

IVESSA, A. S., J.-Q. ZHOU, V. P. SCHULZ, E. K. MONSON and V. A. ZAKIAN, 2002 Saccharomyces Rrm3p, a 5′ to 3′ DNA helicase that promotes replication fork progression through telomeric and subtelomeric DNA. Genes & Development **16:** 1383-1396.

JAGADEESWARAN, P., B. G. FORGET and S. M. WEISSMAN, 1981 Short interspersed repetitive DNA elements in eucaryotes: Transposable DNA elements generated by reverse transcription of RNA pol III transcripts? Cell **26:** 141-142.

JAKUBCZAK, J. L., Y. XIONG and T. H. EICKBUSH, 1990 Type I (R1) and type II (R2) ribosomal DNA insertions of *Drosophila melanogaster* are retrotransposable elements closely related to those of *Bombyx mori*. Journal of Molecular Biology **212:** 37-52.

JANETZKY, B., and L. LEHLE, 1992 Ty4, a new retrotransposon from *Saccharomyces cerevisiae*, flanked by tau-elements. Journal of Biological Chemistry **267:** 19798-19805.

JENSEN, S., M. GASSAMA and T. HEIDMANN, 1999 Taming of transposable elements by homology-dependent gene silencing. Nature Genetics **21:** 209-212.

JIA, Y., S. A. MCADAMS, G. T. BRYAN, H. P. HERSHEY and B. VALENT, 2000 Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. The EMBO Journal **19:** 4004-4014.

JOHNSON, L., 2007 The genome strikes back: the evolutionary importance of defence against mobile elements. Evolutionary Biology **34:** 121-129.

JONES, T., N. A. FEDERSPIEL, H. CHIBANA, J. DUNGAN, S. KALMAN *et al.*, 2004 The diploid genome sequence of *Candida albicans*. Proceedings of the National Academy of Sciences of the United States of America **101:** 7329-7334.

KACHROO, P., M. AHUJA, S. A. LEONG and B. B. CHATTOO, 1997 Organisation and molecular analysis of repeated DNA sequences in the rice blast fungus *Magnaporthe grisea*. Current Genetics **31:** 361-369.

KACHROO, P., S. A. LEONG and B. B. CHATTOO, 1994 Pot2, and inverted repeat transposon from the rice blast fungus *Magnaporthe grisea*. Molecular and General Genetics **245:** 339-348.

KACHROO, P., S. A. LEONG and B. B. CHATTOO, 1995 Mg-SINE: A short interspersed nuclear element from the rice blast fungus *Magnaporthe grisea*. Proceedings of the National Academy of Sciences of the United States of America **92:** 11125-11129.

KANG, S., 2001 Organization and distribution pattern of MGLR-3, a novel retrotransposon in the rice blast fungus *Magnaporthe grisea*. Fungal Genetics and Biology **32:** 11-19.

KANG, S., F. G. CHUMLEY and B. VALENT, 1994 Isolation of the mating-type genes of the pytopathogenic fungus *Magnaporthe grisea* using genomic subtraction. Genetics **138:** 289-296.

KANG, S., M. H. LEBRUN, L. FARRALL and B. VALENT, 2001 Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. Molecular Plant-Microbe Interactions **14:** 671-674.

KANG, S., and L. YONG-HWAN, 2000 Population structure and race variation of the rice blast fungus. The Plant Pathology Journal **16:** 1-8.

KATO, H., M. YAMAMOTO, T. YAMAGUCHI-OZAKI, H. KADOUCHI, Y. IWAMOTO *et al.*, 2000 Pathogenicity, mating ability and DNA restriction fragment length polymorphisms of Pyricularia populations isolated from Gramineae, Bambusideae, and Zingiberaceae plants. Japanese Journal of General Plant Pathology **66:** 30-47.

KATO, S., K. MATSUO, N. NISHIMURA, N. TAKAHASHI and T. TAKANO, 1987 The entire nucleotide-sequence of baboon endogenous virus-DNA - a chimeric genome structure of murine type-C and simian type-D retroviruses. Japanese Journal of Genetics **62:** 127-137.

KEMPKEN, F., and U. KUCK, 1998 Transposons in filamentous fungi-facts and perspectives. Bioessays **20:** 652-659.

KERSHAW, M. J., G. WAKLEY and N. J. TALBOT, 1998 Complementation of the Mpg1 mutant phenotype in *Magnaporthe grisea* reveals functional relationships between fungal hydrophobins. The EMBO Journal **17:** 3838-3849.

KHAN, H., A. SMIT and S. BOISSINOT, 2006 Molecular evolution and tempo of amplification of human LINE-1 retrotransposon since the origin of primates. Genome Research **16:** 78-87.

KIDWELL, M. G., 1992 Horizontal transfer. Current Opinion Genetic Development **2:** 868-873.

KIJIMA, T. E., and H. INNAN, 2010 On the estimation of the insertion time of LTR retrotransposable events. Molecular Biology and Evolution **27:** 896-904.

KILBURN, A. E., M. J. SHEA, R. G. SARGENT and J. H. WILSON, 2001 Insertion of a telomere repeat sequence into a mammalian gene causes chromosome instability. Molecular Cell Biology **21**.

KITO, H., H. TAKAHASHI, J. SATO, S. FUKIYA, T. SONE *et al.*, 2003 Occan, a novel transposon in the *Fot*1 family, is ubiquitously found in several *Magnaporthe grisea* isolates. Current Genetics **42:** 322-331.

KIYOMASU, T., T. MIYAZAWA, T. FURUYA, R. SHIBATA, H. SAKAI *et al.*, 1991 Identification of feline immunodeficiency virus Rev gene activity. Journal of Virology **65:** 4539-4542.

KOBAYASHI, D. Y., S. J. TAMAKI and N. T. KEEN, 1989 Cloned avirulence genes from the tomato pathogen *Pseudomonas syringae* pv. tomato confer cultivar specificity on soybean. Proceedings of the National Academy of Sciences of the United States of America **86:** 157-161.

KOJIMA, K. K., and H. FUJIWARA, 2004 Cross-genome screening of novel sequence-specific non-LTR retrotransposons: various multicopy RNA genes and microsatellites are selected as targets. Molecular Biology and Evolution **21:** 207-217.

KOLOSHA, V. O., and S. L. MARTIN, 2003 High-affinity, non-sequence-specific RNA binding by the open reading frame 1 (ORF1) protein from long interspersed nuclear element 1 (LINE-1). Journal of Biological Chemistry **278:** 8112-8117.

KONIECZNY, A., D. F. VOYTAS, M. P. CUMMINGS and F. M. AUSUBEL, 1991 A superfamily of *Arabidopsis thaliana* retrotransposons. Genetics **127:** 801-809.

KRAUSKOPF, A., and E. H. BLACKBURN, 1996 Control of telomere growth by interactions of RAP1 with the most distal telomeric repeats. Nature **383:** 354-357.

KUCK, U., 1989 The intron of a plastid gene from a green alga contains an open reading frame for a reverse transcriptase-like enzyme. Molecular and General Genetics **218:** 257-265.

KUFF, E. L., A. FEENSTRA, K. LUERDERS, L. SMITH, R. HAWLEY *et al.*, 1983 Intracistemal A-particle genes as moveable elements in the mouse genome. Proceedings of the National Academy of Sciences of the United States of America **80:** 1992-1996.

KUMAR, J., R. J. NELSON and R. S. ZEIGLER, 1999 Population structure and dynamics of *Magnaporthe grisea* in the Indian Himalayas. Genetics **152:** 971-984.

KUPIEC, J. J., A. KAY, M. HAYAT, R. RAVIER, J. PERIES *et al.*, 1991 Sequence-analysis of the simian foamy virus type-1 genome. Gene **101:** 185-194.

KUPIEC, M., and T. D. PETES, 1988 Meiotic recombination between repeated transposable elements in *Saccharomyces cerevisiae*. Molecular and Cellular Biology **8:** 2942-2954.

KUSABA, M., K. ETO, L. D. DON, N. NISHIMOTO, Y. TOSA *et al.*, 1999 Genetic diversity in Pyricularia isolates from various hosts revealed by polymorphisms of nuclear ribosomal DNA and the distribution of the MAGGY retrotransposon. Annual Phytopathological Society of Japan **65:** 588-596.

KUSABA, M., K. HIRATA, Y. SUMIDA, A. YAMAGASHIRA, H. KONGAI-URATA *et al.*, 2006 Molecular genetic characterization and host specificity of Pyrcularia isolates from annual ryegrass in Japan. Plant Pathology Journal **5:** 72-76.

LABBÉ, S., X. YUANNAN and P. H. ROY, 1988 BspMII and AccIII are and isoschizomer pair which differ in their sensitivity to cytosine methylation. Nucleic Acids Research **16:** 7184.

LAMPSON, B. C., J. SUN, M. Y. HSU, J. VALLEJO-RAMIREZ, S. INOUYE *et al.*, 1989 Reverse transcriptase in a clinical strain of *Escherichia coli*: production of branched RNA-linked msDNA. Science **243:** 1033-1038.

LANDSCHOOT, P. J., and B. F. HOYLAND, 1992 Gray leaf-spot of perennial ryegrass turf in Pennsylvania. Plant Disease **76:** 1280-1282.

LANGE, T. D., L. SHIUE, R. M. MYERS, D. R. COX, S. L. NAYLOR *et al.*, 1990 Structure and variability of human chromosome ends. Molecular and Cellular Biology **1990:** 518-527.

LASSMANN, T., and E. L. L. SONNHAMMER, 2005 Kalign - an accurate and fast multiple sequence alignment algorithm. BMC Bioinformatics **6:** 298.

LE ROUZIC, A., and P. CAPY, 2005 The first steps of transposable elements invasion: parasitic strategy vs. genetic drift. Genetics **169:** 1033-1043.

LEVIS, R. W., R. GANESAN, K. HOUTCHENS, L. A. TOLAR and F. M. SHEEN, 1993 Transposons in place of telomeric repeats at a Drosophila telomere. Cell **75:** 1083-1093.

LEVY, M., J. ROMAO, M. A. MARCHETTI and J. E. HAMER, 1991 DNA fingerprinting with a dispersed repeated sequence resolves pathotype diversity in the rice blast fungus. Plant Cell **3:** 95-102.

LI, W., B. WANG, J. WU, G. LU, Y. HU *et al.*, 2009 The *Magnaporthe oryzae* avirulence gene AvrPiz-t encodes a predicted secreted protein that triggers the immunity in rice mediated by the blast resistance gene Piz-t. Molecular Plant-Microbe Interactions **22:** 411-420.

LIM, D., 1992 Structure and biosynthesis of unbranched multicopy single-stranded DNA by reverse transcriptase in a clinical *Escherichia coli* isolate. Molecular Microbiology **6:** 3531-3542.

LIM, D., and W. K. MAAS, 1989 Reverse transcriptase-dependent synthesis of a covalently linked, branched DNA-RNA compound in *E. coli* B. Cell **56:** 891-904.

LINGNER, J., T. R. HUGHES, A. SHEVCHENKO, M. MANN, V. LUNDBLAD *et al.*, 1997 Reverse transcriptase motifs in the catalytic subunit of telomerase. Science **276:** 561-567.

LISKAY, R. M., A. LETSOU and J. L. STACHELEK, 1987 Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cells Genetics **115:** 161-167.

LLACA, V., A. LOU, S. YOUNG and J. MESSING, 1998 Retrotransposable elements of *Sorghum bicolor* (Accession no. AAD27571.1). GenBank Direct Submission.

LLOYD, J. A., A. N. LAMB and S. POTTER, 1987 Phylogenetic screening of the human genome: identification of differentially hybridizing repetitive sequence families. Molecular Biology and Evolution **4:** 85-98.

LOAYZA, D., and T. DE LANGE, 2003 POT1 as a terminal transducer of TRF1 telomere length control. Nature **423:** 1013-1018.

LOEB, D. D., R. W. PADGETT, S. C. HARDIES, W. R. SHEHEE, M. B. COMER *et al.*, 1986 The sequence of a large L1Md element reveals a tandemly repeated 5' end and several features found in retrotransposons. Molecular and Cellular Biology **6:** 168-182.

LOHE, A., and D. HARTL, 1996 Autoregulation of mariner transposase activity by overproduction and dominant-negative complementation. Molecular Biology and Evolution **13:** 549-555.

LOUIS, E. J., 1995 The chromosome ends of *Saccharomyces cerevisiae*. Yeast **11:** 1553-1573.

LOUIS, E. J., and J. E. HABER, 1990 Mitotic recombination among subtelomeric Y' repeats in *Saccharomyces cerevisiae*. Genetics **124:** 547-559.

LUAN, D. D., and T. H. EICKBUSH, 1995 RNA template requirements for target DNA-primed reverse transcritpion by R2 Transposable element. Cell Biology **15:** 3882-3891.

LUAN, D. D., and T. H. EICKBUSH, 1996 Downstream 28S gene sequences on the RNA template affect the choice of primer and the accuracy of initiation by the R2 reverse transcriptase. Molecular Cell Biology **16:** 4726–4734.

LUAN, D. D., M. H. KORMAN, J. L. JAKUBCZAK and T. H. EICKBUSH, 1993 Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell **72:** 595-605.

MA, L.-J., H. C. VAN DER DOES, K. A. BORKOVICH, J. J. COLEMAN, M.-J. DABOUSSI *et al.*, 2010 Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. Nature **464:** 367-373.

MACKILL, A. O., and J. M. BONMAN, 1986 New hosts of *Pyricularia oryzae*. Plant Disease **70:** 125-127.

MAKINO, K., K. ISHII, T. YASUNAGA, M. HATTORI, K. YOKOYAMA *et al.*, 1998 Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic *Escherichia coli* O157:H7 derived from Sakai outbreak. DNA Research **5:** 1-9.

MAKOVETS, S., I. HERSKOWITZ and E. H. BLACKBURN, 2004 Anatomy and dynamics of DNA replication fork movement in yeast telomeric regions. Molecular and Cellular Biology **24:** 4019-4031.

MALCA, I., and J. H. OWEN, 1957 The gray leaf spot disease of St. Augustinegrass. Plant Disease Report **41:** 871-875.

MALIK, H. S., W. D. BURKE and T. H. EICKBUSH, 1999 The age and evolution of non-LTR retrotransposable elements. Molecular Biology and Evolution **16:** 793-805.

MALIK, H. S., and T. H. EICKBUSH, 1999 Retrotransposable elements R1 and R2 in the rDNA units of *Drosophila mercatorum*: abnormal abdomen revisited. Genetics **151:** 653-665.

MALIK, H. S., and T. H. EICKBUSH, 2001 Phylogenetic analysis of Ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses. Genome Research **11:** 1187-1197.

MALIK, H. S., S. HENIKOFF and T. H. EICKBUSH, 2000 Poised for contagion: evolutionary origins of the infectious abilites of the invertebrate retroviruses. Genome Research **10:** 1307-1318.

MARLOR, R. L., S. M. PARKHURST and V. G. CORCES, 1986 The *Drosophila melanogaster* gypsy transposable element encodes putative gene products homologous to retroviral proteins. Molecular and Cellular Biology **6:** 1129-1134.

MARTIN, S. L., and F. D. BUSHMAN, 2001 Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. Molecular Cell Biology **21:** 467-475.

MARVIN, M. E., C. D. GRIFFIN, D. E. EYRE, D. B. BARTON and E. J. LOUIS, 2009 In *Saccharomyces cerevisiae*, yKU and subtelomeric core X sequences repress

homologous recombination near telomeres as part of the same pathway. Genetics **183:** 441-451.

MASON, J. M., and H. BIESSMANN, 1995 The unusual telomeres of Drosophila. Trends in Genetics **11:** 58-62.

MCCLELLAND, M., M. NELSON and E. RASCHKE, 1994 Effect of site-specific modification on restriction endonucleases and DNA modification methyltransferases. Nucleic Acids Research **22:** 3640-3659.

MCCLINTOCK, B., 1939 The behavior in successive nuclear divisions of a chromosome broken at meiosis. Proceedings of the National Academy of Sciences of the United States of America **25:** 405-416.

MCCLINTOCK, B., 1941 The stability of broken ends of chromosomes in *Zea mays*. Genetics **26:** 234-282.

MCEACHERN, M. J., D. H. UNDERWOOD and E. H. BLACKBURN, 2002 Dynamics of telomeric DNA turnover in yeast. Genetics **160:** 63-73.

MEFFORD, H. C., and B. J. TRASK, 2002 The complex structure and dynamic evolution of human subtelomeres. Nature Reviews Genetics **3:** 91-102.

MENON, S. K., B. J. EILERS, M. J. YOUNG and C. M. LAWRENCE, 2010 The crystal structure of D212 from sulfolobus spindle-shaped virus Ragged Hills reveals a new member of the PD-(D/E)XK nuclease superfamily. Journal of Virology **84:** 5890-5897.

MEYER, D. H., and A. M. BAILIS, 2008 Telomerase deficiency affects the formation of chromosomal translocations by homologous recombination in *Saccharomyces cerevisiae*. PLoS ONE **3:** e3318.

MEYN, M. A., L. FARRAL, B. VALENT, F. G. CHUMLEY and M. J. ORBACH, *Magnaporthe grisea* repeated DNA element MGR583 is a member of the LINE-1 class of polyA retrotransposons, pp.

MIAO, V., S. COVERT and H. VANETTEN, 1991 A fungal gene for antibiotic resistance on a dispensable (“B”) chromosome. Science **254:** 1773-1776.

MICHELS, C. A., and R. B. NEEDLEMAN, 1984 The dispersed, repeated family of MAL loci in Saccharomyces spp. Journal of Bacteriology **157:** 949-952.

MILGROOM, M. G., 1996 Recombination and the multilocus structure of fungal populations. Annual Review of Phytopathology **34:** 457-477.

MIZROKHI, L. J., S. G. GEORGIEVA and Y. V. ILYAN, 1988 Jockey, a mobile Drosophila element similar to mammalian LINEs, is transcribed from the internal promoter by RNA polymerase II. Cell **54:** 685-691.

MIZROKHI, L. J., and A. M. MAZO, 1990 Evidence for horizontal transmission of the mobile element jockey between distant Drosophila species. Proceedings of the National Academy of Sciences of the United States of America **87:** 9216-9220.

MORRISON, H. G., A. G. MCARTHUR, F. D. GILLIN, S. B. ALEY, R. D. ADAM *et al.*, 2007 Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. Science **317:** 1921-1926.

MOUCHES, C., N. BENSAADI and J. C. SALVADO, 1992 Characterization of a LINE retroposon dispersed in the genome of three non-sibling Aedes mosquito species. Gene **120:** 183-190.

MOUNT, S. M., and G. M. RUBIN, 1985 Complete nucleotide sequence of the Drosophila transposable element copia: homology between copia and retroviral proteins. Molecular and Cellular Biology **5:** 1630-1638.

MURAKAMI, J., V. T. B. HAU, S. MAYAMA and Y. TOSA, 2006 Reduction in aggressiveness among hybrids between host-specific pathotypes of *Magnaporthe oryzae* is caused by reduced ability ot overcome adult resistance at the level of penetration. Journal of General Plant Pathology **72:** 284-291.

MUSIO, A., G. RAINALDI and I. SBRANA, 1996 Spontaneous and aphidicolinsensitive fragile site 3cen co-localizes with the (TTAGGG)n telomeric sequence in Chinese hamster cells. Cytogenetics Cell Genetics **75:** 159-163.

NAKAMURA, T. M., G. B. MORIN, K. B. CHAPMAN, S. L. WEINRICH, W. H. ANDREWS *et al.*, 1997 Telomerase catalytic subunit homologs from fission yeast and human. Science **277:** 955-959.

NAKAYASHIKI, H., K. IKEDA, Y. HASHIMOTO, Y. TOSA and S. MAYAMA, 2001a Methylation is not the main force repressing the retrotransposon MAGGY in *Magnaporthe grisea*. Nucleic Acids Research **29:** 1278-1284.

NAKAYASHIKI, H., H. MATSUO, I. CHUMA, K. IKEDA, S. BETSUYAKU *et al.*, 2001b Pyret, a Ty3/Gypsy retrotransposon in *Magnaporthe grisea* contains an extra domain between the nucleocapsid and protease domains. Nucleic Acids Research **29:** 4106-4113.

NATARAJAN, S., C. GROFF-VINDMAN and M. J. MCEACHERN, 2003 Factors influencing the recombinational expansion and spread of telomeric tandem arrays in *Kluyveromyces lactis*. Eukaryotic Cell **2:** 1115-1127.

NATARAJAN, S., and M. J. MCEACHERN, 2002 Recombinational telomere elongation promoted by DNA circles. Molecular Cell Biology **22:** 4512-4521.

NAUMOV, G., H. TURAKAINEN, E. NAUMOVA, S. AHO and M. KORHOLA, 1990 A new family of polymorphic genes in *Saccharomyces cerevisiae*: Alpha-galactosidase genes MEL1-MEL7. Molecular and General Genetics **224:** 119-128.

NISHIMURA, M., N. HAYASHI, N. S. JWA, G. W. LAU, J. E. HAMER *et al.*, 2000 Insertion of the LINE retrotransposon MGL causes a conidiophore pattern mutation in *Magnaporthe grisea*. Molecular Plant-Microbe Interactions **13:** 892-894.

NITTA, N., M. L. FARMAN and S. A. LEONG, 1997 Genome organization of Magnaporthe grisea: integration of genetic maps, clustering of transposable elements and identification of genome duplications and rearrangements. Theoretical and Applied Genetics **95:** 20-32.

NOGARE, D. E. D., M. S. CLARK, G. ELGAR, I. G. FRAME and R. T. M. POULTER, 2002 Xena, a full-length basal retroelement from tetraodontid fish. Molecular Biology and Evolution **19:** 247-255.

NOGUCHI, M. T., N. YASUDA and Y. FUJITA, 2006 Evidence of genetic exchange by parasexual recombination and genetic analysis of pathogenicity and mating type of parasexual recombinants in rice blast bungus, *Magnaporthe oryzae*. Phytopathology **96:** 746-750.

NOVIKOVA, O., and A. BLINOV, 2005 NeSL-1 like element from *Caenorhabditis briggsae*. (Accession no. AAZ15238.1), pp. in *GenBank Direct Submission*.

NOVIKOVA, O., V. FET and A. BLINOV, 2009 Non-LTR retrotransposons in fungi. Functional & Integrative Genomics **9:** 27-42.

NOVIKOVA, O., J. STARNES, D. THORNBURY and M. FARMAN, 2011 Elucidating mechanisms of telomere instability in *Magnaporthe oryzae*: retrotransposons running riot. Fungal Genetics Reports **58-Supplement:** 263.

O'HARE, K., M. R. ALLEY, T. E. CULLINGFORD, A. DRIVER and M. J. SANDERSON, 1991 DNA sequence of the Doc retroposon in the white-one mutant of *Drosophila melanogaster* and of secondary insertions in the phenotypically altered derivatives white-honey and white-eosin. Molecular and General Genetics **225:** 17-24.

ODA, K., K. YAMATO, E. OHTA, Y. NAKAMURA, M. TAKEMURA *et al.*, 1992 Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. Journal of Molecular Biology **223:** 1-7.

OKADA, N., M. HAMADA, I. OGIWARA and K. OHSHIMA, 1997 SINEs and LINEs share common 3' sequences: a review. Gene **205:** 229-243.

OKAMOTO, H., M. IMAI, M. SHIMOZAKI, Y. HOSHI, H. IIZUKA *et al.*, 1986 Nucleotide sequence of a cloned hepatitis B virus genome, subtype ayr: comparison with genomes of the other three subtypes. Journal of General Virology **67:** 2305-2314.

OKAZAKI, S., H. ISHIKAWA and H. FUJIWARA, 1995 Structural analysis of TRAS1, a novel family of telomeric repeat-associated retrotransposons in the silkworm, *Bombyx mori*. Molecular and Cellular Biology **15:** 4545-4552.

ORBACH, M. J., L. FARRALL, J. A. SWEIGARD, F. G. CHUMLEY and B. VALENT, 2000 A telomeric avirulence gene determines efficacy for the rice blast resistance gene Pi-ta. Plant Cell **12:** 2019-2032.

OSANAI, M., K. K. KOJIMA, R. FUTAHASHI, S. YAGUCHI and H. FUJIWARA, 2006 Identification and characterization of the telomerase reverse transcriptase of *Bombyx mori* (silkworm) and *Tribolium castaneum* (flour bettle). Gene **376:** 281-289.

OSANAI, M., H. TAKAHASHI, K. K. KOJIMA, M. HAMADA and H. FUJIWARA, 2004 Essential motifs in the 3' untranslated region required for retrotransposition and the precise start of reverse transcritiption in the non-long-terminal-repeat retrotransposon SART1. Molecular Cell Biology **24:** 7902-7913.

OU, S. H., 1980 A look a worldwide rice blast control. Plant Disease **64:** 439-445.

PALM, W., and T. DE LANGE, 2008 How shelterin protects mamalian telomeres. Annual Review of Genetics **42:** 301-334.

PARDUE, M. L., and P. G. DEBARYSHE, 2003 Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. Annual Review of Genetics **37:** 485-511.

PARK, C.-H., S. CHEN, G. SHIRSEKAR, B. ZHOU, C. H. KHANG *et al.*, 2012 The *Magnaporthe oryzae* effector AvrPiz-t targets the RING E3 ubiquitin ligase APIP6 to suppress pathogen-associated molecular pattern–triggered immunity in rice. The Plant Cell**:** 1-15.

PARK, S.-Y., M.-H. CHI, M. G. MILGROOM, H. KIM and S.-S. HAN, 2010 Genetic Stability of Magnaporthe oryzae during Successive Passages through Rice Plants and on Artificial Media. The Plant Pathology Journal **26:** 313-320.

PARK, S. Y., M. G. MILGROOM, S. S. HAN, S. KANG and Y. H. LEE, 2008 Genetic differentiation of *Magnaporthe oryzae* populations from scouting plots and commercial rice fields in Korea. Phytopathology **98:** 436-442.

PARLEVLIET, J., 2002 Durability of resistance against fungal, bacterial, and viral pathogen; present situation. Euphytica **124:** 147-156.

PEARCE, S. R., U. PICH, G. HARRISON, A. J. FLAVELL, J. S. HESLOP-HARRISON *et al.*, 1996 The Ty1-copia group retrotransposons of *Allium cepa* are distributed throughout the chromosomes but are enriched in the terminal heterochromatin. Chromosome Research **5:** 357-364.

PETRZIK, K., V. BENES, I. MRAZ, J. HONETSLEGROVA-FRANOVA, W. ANSORGE *et al.*, 1998 Strawberry vein banding virus--definitive member of the genus caulimovirus. Virus Genes **16:** 303-305.

PEYYALA, R., and M. L. FARMAN, 2006 *Magnaporthe oryzae* isolates causing gray leaf spot of perennial ryegrass possess a functional copy of the AVR1-CO39 avirulence gene. Molecular Plant Pathology **7:** 157-165.

PLASTERK, R. H., and J. T. GROENEN, 1992 Targeted alterations of the *Caenorhabditis elegans* genome by transgene instructed DNA double strand break repair following Tc1 excision. The EMBO Journal **11:** 287-290.

POLAND, J. A., P. J. BALINT-KURTI, R. J. WISSER, R. C. PRATT and R. J. NELSON, 2008 Shades of gray: the world of quantitative disease resistance. Trends in Plant Science **14:** 21-29.

POWER, M. D., P. A. MARX, M. L. BRYANT, M. B. GARDNER, P. J. BARR *et al.*, 1986 Nucleotide-sequence of SRV-1, a type-D simian acquired-immune-deficiency-syndrome retrovirus. Science **231:** 1567-1572.

PRABHU, A., H. G. MORRISON, C. R. MARTINEZ, 3RD and R. D. ADAM, 2007 Characterisation of the subtelomeric regions of *Giardia lamblia* genome isolate WBC6. International Journal for Parasitology **37:** 503-513.

PRUD'HOMME, N., M. GANS, M. MASSON, C. TERZIAN and A. BUCHETON, 1995 Flamenco, a gene controlling the gypsy retrovirus of *Drosophila melanogaster*. Genetics **139:** 606-611.

RACHIDI, N., M. J. MARTINEZ, P. BARRE and B. BLONDIN, 2000 *Saccharomyces cerevisiae* PAU genes are induced by anaerobiosis. Molecular Microbiology **35:** 1421-1430.

REDDY, A. P. K., and J. M. BONMAN, 1987 Recent epidemics of Rice Blast in India and Egypt. Plant Disease **71:** 850.

REHMEYER, C., W. LI, M. KUSABA, Y.-S. KIM, D. BROWN *et al.*, 2006 Organization of chromosome ends in the rice blast fungus, *Magnaporthe oryzae*. Nucleic Acids Research **34:** 4685-4701.

RENSHAW, R. W., and J. W. CASEY, 1994 Transcriptional mapping of the 3' end of the bovine syncytial virus genome. Journal of Virology **68:** 1021-1028.

RICCHETTI, M., B. DUJON and C. FAIRHEAD, 2003 Distance from the chromsome end determines the efficiency of double strand break repair in subtelomeres of haploid yeast. Journal of Molecular Biology **328:** 847-862.

RICHINS, R. D., 1993 Organization and transcription of the peanut chlorotic streak caulimovirus genome (Accession no. NP_042513.1). GenBank Direct Submission.

RICHINS, R. D., H. B. SCHOLTHOF and R. J. SHEPHERD, 1987 Sequence of figwort mosaic virus DNA (caulimovirus group). Nucleic Acids Research **15:** 8451-8466.

ROSSMAN, A. Y., R. J. HOWARD and B. VALENT, 1990 *Pyricularia grisea*, the correct name for the rice blast disease fungus. Mycologia **82:** 509-512.

ROUMEN, E., M. LEVY and J. L. NOTTEGHEM, 1997 Characterisation of the European pathogen population of *Magnaporthe grisea* by DNA fingerprinting and pathotype analysis. European Journal of Plant Pathology **103:** 363-371.

RUBNITZ, J., and S. SUBRAMANI, 1984 The minimum amount of homology required for homologous recombination in mammalian cells. Molecular and Cellular Biology **4:** 2253-2258.

RUDD, M. K., C. FRIEDMAN, S. S. PARGHI, E. V. LINARDOPOULOU, L. HSU *et al.*, 2007 Elevated rates of sister chromatid exchange at chromosome ends. PLoS Genet **3:** e32.

SAIGO, K., W. KUGIMIYA, Y. MATSUO, S. INOUYE, K. YOSHIOKA *et al.*, 1984 Identification of the coding sequence for a reverse transcriptase-like enzyme in a transposable genetic element in *Drosophila melanogaster*. Nature **312:** 659-661.

SAINTIGNY, Y., F. DELACÔTE, G. VARÈS, F. PETITOT, S. LAMBERT *et al.*, 2001 Characterization of homologous recombination induced by replication inhibition in mammalian cells. The EMBO Journal **20:** 3861-3870.

SAITOU, N., and M. NEI, 1987 The neighbor-joining method - a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution **4:** 406-425.

SALEH, D., J. MILAZZO, H. ADREIT, D. THARREAU and E. FOURNIER, 2012 Asexual reproduction induces a rapid and permanent loss of sexual reproduction capacity in the rice fungal pathogen, *Magnaporthe oryzae*: results of in vitro experimental evolution assays. BMC Evolutionary Biology **12:** 42.

SALTARELLI, M., G. QUERAT, D. A. M. KONINGS, R. VIGNE and J. E. CLEMENTS, 1990 Nucleotide-sequence and transcriptional analysis of molecular clones of CAEV Which generate infectious virus. Virology **179:** 347-364.

SAMPATHI, S., and W. CHAI, 2011 Telomere replication: poised but puzzling. Journal of Cellular and Molecular Medicine **15:** 3-13.

SANCHEZ, E., K. ASANO and T. SONE, 2011 Characterization of *Inago*1 and *Inago*2 retrotransposons in *Magnaporthe oryzae*. Journal of General Plant Pathology**:** 1-4.

SANDELL, L. L., and V. A. ZAKAIN, 1993 Loss of a yeast telomere: arrest, recovery, and chromosome loss. Cell **75:** 729-739.

SANZ-ALFEREZ, S., P. SANMIGUEL, Y. K. JIN, P. S. SPRINGER and J. L. BENNETZEN, 2003 Structure and evolution of the *Cinful* retrotransposon family of maize. Genome **46:** 745-752.

SASSAMAN, D. M., B. A. DOMBROSKI, J. V. MORAN, M. L. KIMBERLAND, T. P. NAAS *et al.*, 1997 Many human L1 elements are capable of retrotransposition. Nature Genetics **16:** 37-43.

SAVITSKY, M., D. KWON, P. GEORGIEV, A. KALMYKOVA and V. GVOZDEV, 2006 Telomere elongation is under the control of the RNAi-based mechanism in the Drosophila germline. Genes & Development **20:** 345-354.

SCHÜRCH, S., C. C. LINDE, W. KNOGGE, L. F. JACKSON and B. A. MCDONALD, 2004 Molecular population genetic analysis differentiates two virulence mechanisms of the fungal avirulence gene NIP1. Molecular Plant-Microbe Interactions **10:** 1114-1125.

SCHWARTZ, D. E., R. TIZARD and W. GILBERT, 1983 Nucleotide-sequence of rous-sarcoma virus. Cell **32:** 853-869.

SEIKI, M., S. HATTORI, Y. HIRAYAMA and M. YOSHIDA, 1983 Human adult T-cell leukemia-virus - complete nucleotide-sequence of the provirus genome integrated in leukemia-cell DNA. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences **80:** 3618-3622.

SELENT, U., T. RUTER, E. KOHLER, M. LIEDTKE, V. THIELKING *et al.*, 1992 A site-directed mutagenesis study to identify amino acid residues involved in the catalytic function of the restriction endonuclease *Eco*RV. Biochemstry **31:** 4808-4815.

SELKER, E. U., E. B. CAMBARERI, B. C. JENSEN and K. R. HAACK, 1987 Rearrangement of duplicated DNA in specialized cells of Neurospora. Cell **51:** 741-752.

SHAMPAY, J., and E. H. BLACKBURN, 1988 Generation of telomere-length heterogeneity in Saccharomyces cervisiae. Genetics **85:** 534-538.

SHEEN, F. M., and R. W. LEVIS, 1994 Transposition of the LINE-like retrotransposon TART to Drosophila chromosome termini. Proceedings of the National Academy of Sciences of the United States of America **91:** 12510-12514.

SHEN, W. C., P. BOBROWICZ and D. J. EBBOLE, 1999 Isolation of pheromone precursor genes of *Magnaporthe grisea*. Fungal Genetics and Biology **27:** 253-263.

SILVA, J. C., E. L. LORETO and J. B. CLARK, 2004 Factors that affect the horizontal transfer of transposable elements. Current Issues Molecular Biology **6:** 57-71.

SKAMNIOTI, P., and S. J. GURR, 2007 *Magnaporthe grisea* Cutinase2 mediates appressorium differentiation and host penetration and is required for full virulence. The Plant Cell **19:** 2674-2689.

STAGE, D. E., and T. H. EICKBUSH, 2009 Origin of nascent lineages and the mechanisms used to prime second-strand DNA synthesis in the R1 and R2 retotransposons of *Drosophila*. Genome Biology **10:** R49.

STARNES, J. H., D. H. THORNBURY, C. J. REHMEYER, O. S. NOVIKOVA and M. FARMAN, 2012 Telomere-targeted retrotransposons in the rice blast fungus *Magnaporthe oryzae*: agents of telomere instability. Genetics **191:** 389-406.

STERGIOPOULOS, I., and P. J. G. M. DE WIT, 2009 Fungal Effector Proteins. Annual Review of Phytopathology **47:** 233-263.

STRUMBERG, D., A. A. PILON, M. SMITH, R. HICKEY, L. MALKAS *et al.*, 2000 Conversion of topoisomerase I cleavage complexes on the leading strand of ribosomal DNA into 5'-phosphorylated DNA double-strand breaks by replication runoff. Cell Biology**:** 3977-3987.

SUZUKI, H., 1975 Meterological factors in the epidemiology of rice blast. Annual Review of Phytopathology **13:** 239-256.

SWEIGARD, J. A., A. M. CARROLL, S. KANG, L. FARRALL, F. G. CHUMLEY *et al.*, 1995 Identification, cloning, and characterization of *Pwl2*, a gene for host species-specificity in the rice blast fungus. Plant Cell **7:** 1221-1233.

SZAK, S. T., O. K. PICKERAL, W. MAKALOWSKI, M. S. BOGUSKI, D. LANDSMAN *et al.*, 2002 Molecular archeology of L1 insertions in the human genome. Genome Biology **3:** research0052-research0052.0018.

SZOSTAK, J. W., and R. WU, 1980 Unequal crossing over in the ribosomal DNA of *Saccharomyces cerevisiae*. Nature **284:** 426-430.

TAKAHASHI, H., S. OKAZAKI and H. FUJIWARA, 1997 A new family of site-specific retrotransposons, SART1, is inserted into telomeric repeats of the silkworm, *Bombyx mori*. Nucleic Acids Research **25:** 1578-1584.

TALBOT, N. J., and A. J. FOSTER, 2001 Genetics and genomics of the rice blast fungus *Magnaporthe grisea*: developing an experimental model for understanding fungal diseases of cereals. Advances in Botanical Research Incorporating Advances in Plant Pathology **34:** 263-287.

TAMURA, K., J. DUDLEY, M. NEI and S. KUMAR, 2007 MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution **24:** 1596-1599.

TANAKA, M., and H. NAKAYASHIKI, 2009 Population structure of *Eleusine* isolates of *Pyricularia oryzae* and its evolutionary relationships. Japanese Journal of General Plant Pathology **75:** 173-180.

TAVAKOLI, N. P., and K. M. DERBYSHIRE, 2001 Tipping the balance between replicative and simple transposition. The EMBO Journal **20:** 2923-2930.

THOMPSON, J. K., J. P. RUBIOA, S. CARUANA, A. BROCKMAN, M. E. WICKHAM *et al.*, 1997 The chromosomal organization of the *Plasmodium falciparum* var gene family is conserved. Molecular and Biochemical Parasitology **87:** 49-60.

THON, M., H. PAN, S. DIENER, J. PAPALAS, A. TARO *et al.*, 2006 The role of transposable element clusters in genome evolution and loss of synteny in the rice blast fungus *Maganporthe oryzae*, pp. R16 in *Genome Biology*.

THON, M. R., S. L. MARTIN, S. GOFF, R. A. WING and R. A. DEAN, 2004 BAC end sequences and a physical map reveal transposable element content and clustering patterns in the genome of *Magnaporthe grisea*. Fungal Genetics and Biology **41:** 657-666.

TONG, S., F. MATTES, K. TEUBNER and H. E. BLUM, 1990 Complete nucleotide sequence of a chinese duck hepatitis B virus. Nucleic Acids Research **18:** 6139.

TOSA, Y., K. HIRATA, H. TAMBA, S. NAKAGAWA, I. CHUMA *et al.*, 2004 Genetic constitution and pathogenicity of Lolium isolates of *Magnaporthe oryzae* in comparison with host species-specific pathotypes of the blast fungus. Phytopathology **94:** 454-462.

TOSA, Y., J. OSUE, Y. ETO, H.-S. OH, H. NAKAYASHIKI *et al.*, 2005 Evolution of an avirulence gene, AVR1-CO39, concmitant with the evolution and differentiation of *Magnaporthe oryzae*. Molecular Plant-Microbe Interactions **18:** 1148-1160.

TOSA, Y., W. UDDIN, G. VIJI, S. KANG and S. MAYAMA, 2007 Comparative genetic analysis of *Magnaporthe oryzae* isolates causing gray leaf spot of perennial ryegrass turf in the United States and Japan. Plant Disease **91:** 517-524.

TREDWAY, L. P., K. L. STEVENSON and L. L. BURPEE, 2005 Genetic structure of *Magnaporthe grisea* populations associated with St. Augustinegrass and tall fescue in Georgia. Phytopathology **95:** 463-471.

TU, Z., J. ISOE and J. A. GUZOVA, 1998 Structural, genomic, and phylogenetic analysis of Lian, a novel family of non-LTR retrotransposons in the yellow fever mosquito, *Aedes aegypti*. Molecular Biology and Evolution **15:** 837-853.

UDDIN, W., G. VIJI and P. VINCELLI, 2003 Gray leaf spot (blast) of perennial ryegrass turf: An emerging problem for the turfgrass industry. Plant Disease **87:** 880-889.

URASHIMA, A. S., Y. HASHIMOTO, L. D. DON, M. KUSABA, Y. TOSA *et al.*, 1999 Molecular analysis of the wheat blast population in brazil with a homolog of retrotransoson of MGR583. Annual Phytopathological Society of Japan **65:** 429-436.

VAN DEN ACKERVEKEN, G. F. J. M., J. A. VAN KAN, M. H. A. J. JOOSTEN, J. M. MUISERS, H. M. VERBAKEL *et al.*, 1993 Characterization of two putative pathogenicity genes of the fungal tomato pathogen *Cladosporium fulvum*. Molecular Plant-Microbe Interactions **6:** 210-215.

VAN STEENSEL, B., and T. DE LANGE, 1997 Control of telomere length by the human telomeric protein TRF1. Nature **385:** 740-743.

VANNIER, J.-B., V. PAVICIC-KALTENBRUNNER, M. I. R. PETACORIN, H. DING and S. J. BOULTON, 2012 RTEL1 dismantles t loops and counteracts telomeric G4-DNA to maintain telomere integreity. Cell **149:** 795-806.

VERSTREPEN, K. J., and G. R. FINK, 2009 Genetic and epigenetic mechanisms underlying cell-surface variablity in protozoa and fungi. Annual Review of Genetics **43:** 1-24.

VERSTREPEN, K. J., and F. M. KLIS, 2006 Flocculation, adhesion and biofilm formation in yeasts. Molecular Microbiology **60:** 5-15.

VIJI, G., S. KANG and W. UDDIN, 2001 *Pyricularia grisea* causing gray leaf spot of perennial ryegrass turf: population structure and host specificity. Plant Disease **85:** 817-826.

VIJI, G., and W. UDDIN, 2002 Distribution of mating type alleles and fertility status of *Magnaporthe grisea* causing gray leaf spot of perennial ryegrass and St. Augustinegrass turf. Plant Disease **86:** 827-832.

VILLANUEVA, M. S., S. P. WILLIAMS, C. B. BEARD, F. F. RICHARDS and S. AKSOY, 1991 A new member of a family of site-specific retrotransposons is present in the spliced leader RNA genes of *Trypanosoma cruzi*. Molecular and Cellular Biology **11:** 6139-6148.

VINCELLI, P., 1999 Gray leaf spot: an emerging disease of perennial ryegrass. Turfgrass Trends **7:** 1-8.

VINCELLI, P., 2000 Fungicidal control of gray leaf spot. Golf Course Management **68:** 54-61.

VINCELLI, P., and E. DIXON, 2002 Update: Fungicide failure against gray leaf spot. Golf Course Management **70:** 53-55.

VOLFF, J.-N., C. KORTING, A. FROSCHAUER, K. SWEENEY and M. SCHARTL, 2001 Non-LTR retrotransposons encoding a restriction enzyme-like endonuclease in vertebrates. Journal of Molecular Evolution **52:** 351-360.

WADA, M., and Y. NAKAMURA, 1996 Antigenic variation by telomeric recombination of major-surface-glycoprotein genes of *Pneumocystis carinii*. Journal of Eukaryotic Microbiology **43:** S8-S8.

WALLACE, M., L. ANDERSON, A. SAULINO, P. GREGORY, T. GLOVER *et al.*, 1991 A *de novo* Alu insertion results in neurofibromatosis Type 1. Nature **353:** 864-866.

WANG, R. C., A. SMOGORZEWSKA and T. D. LANGE, 2004 Homologous recombination generates T-loop-sized deletions at human telomeres. Cell **119:** 355-366.

WATERHOUSE, A. M., J. B. PROCTER, D. M. A. MARTIN, M. CLAMP and G. J. BARTON, 2009 Jalview version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics **25:** 1189-1191.

WICKER, T., F. SABOT, A. HUA-VAN, J. L. BENNETZEN, P. CAPY *et al.*, 2007 A unified classification system for eukaryotic transposable elements. Nature Reviews Genetics **8:** 973-982.

WILHELM, M., F. WILHELM, G. KEITH, B. AGOUTIN and H. T., 1994 Yeast Ty1 retrotransposon: the minus-strand primer binding site and a cis-acting domain of the Ty1 RNA are both important for packaging of primer tRNA inside virus-like particles. Nucleic Acids Reasearch **22:** 4560-4565.

WOUW, A. P. V. D., A. J. COZIJNSEN, J. K. HANE, P. C. BRUNNER, B. A. MCDONALD *et al.*, 2010 Evolution of Linked Avirulence Effectors in *Leptosphaeria maculans* Is Affected by Genomic Environment and Exposure to Resistance Genes in Host Plants. PLoS Pathogens **6:** e1001180.

WU, C., Y.-S. KIM, K. M. SMITH, W. LI, H. M. HOOD *et al.*, 2009 Characterization of chromosome ends in the filamentous fungus *Neurospora crassa*. Genetics **181:** 1129-1145.

XIONG, Y., and T. H. EICKBUSH, 1988 The site-specific ribosomal DNA insertion element R1Bm belongs to a class of non-long-terminal-repeat retrotransposons. Molecular and Cellular Biology **8:** 114-123.

XIONG, Y., and T. H. EICKBUSH, 1990 Origin and evolution of retroelements based upon their reverse transcriptase sequences. The EMBO Journal **9:** 3353-3362.

XIONG, Y., and T. H. EICKBUSH, 1993 Dong, a non-long terminal repeat (non-LTR) retrotransposable element from *Bombyx mori*. Nucleic Acids Research **21:** 1318.

XUE, M., J. YANG, Z. LI, S. HU, N. YAO *et al.*, 2012 Comparative analysis of the genomes of two field isolates of the rice blast fungus *Magnaporthe oryzae*. PLoS Genetics **8:** e1002869.

YAEGASHI, H., and S. UDAGAWA, 1978 The taxonomical identity of the perfect sate of *Pyricularia grisea* and its allies. Canadian Journal of Botany **56:** 180-183.

YAMAGASHIRA, A., C. IWAI, M. MOROISHI, M. MISAKA, Y. FUGITA *et al.*, 2008 Population structure of *Magnaporthe oryzae* isolates from green foxtail in Japan examined by DNA fingerprint analysis. Myoscience **49:** 351-358.

YANG, J., H. S. MALIK and T. H. EICKBUSH, 1999 Identification of the enonuclease domain encoded by R2 and other site-specific, non-long terminal repeate retrotransposable elments. Proceedings of the National Academy of Sciences of the United States of America **96:** 7847-7852.

YUKI, S., S. INOUYE, S. ISHIMARU and K. SAIGO, 1986 Nucleotide sequence characterization of a Drosophila retrotransposon, 412. European Journal of Biochemistry **158:** 403-410.

ZAGROBELNY, M., D. C. JEFFARES and P. ARCTANDER, 2004 Differences in non-LTR retrotransposons within *C. elegans* and *C. briggsae* genomes. Gene **330:** 61-66.

ZAGULSKI, M., B. BABINSKA, R. GROMADKA, A. MIGDALSKI, J. RYTKA *et al.*, 1995 The sequence of 24.3 kb from chromosome X reveals five complete open reading frames, all of which correspond to new genes, and a tandem insertion of a Ty1 transposon. Yeast **11:** 1179-1186.

ZAKIAN, V. A., 1995 Telomeres: beginning to understand the end. Science **270:** 1601-1607.

ZEIGLER, R. S., and F. J. CORREA, 2000 Applying *Magnaporthe grisea* population analysis for durable rice blast resistance, pp. in *APSnet Features*.

ZHENG, Y., W. ZHENG, F. LIN, Y. ZHANG, Y. YI *et al.*, 2011 AVR1-CO39 is a predominant locus governing the broad avirulence of Magnaporthe oryzae 2539 on cultivated rice (Oryza sativa L.). Molecular Plant-Microbe Interactions **24:** 13-17.

ZHOU, E., Y. JIA, P. SINGH, J. C. CORRELL and F. N. LEE, 2007 Instablity of the *Magnaporthe oryzae* avirulence gene *AVR-Pita* alters virulence. Fungal Genetics and Biology **44:** 1024-1034.

ZINGLER, N., O. WEICHENRIEDER and G. SCHUMANN, 2005a APE-type non- LTR retrotransposons: determinants involved in target site recognition. Cytogenetic and Genome Research **110:** 250-268.

ZINGLER, N., U. WILLHOEFT, H.-P. BROSE, V. SCHODER, T. JAHNS *et al.*, 2005b Analysis of 5 junctions of human LINE-1 and Alu retrotransposons suggests an alternative model for 5-end attachment requiring microhomology-mediated end-joining. Genome Research **15:** 780-789.

ZOU, S., D. A. WRIGHT and D. F. VOYTAS, 1995 The Saccharomyces Ty5 retrotransposon family is associated with origins of DNA replication at the telomeres and the silent mating locus HMR. Proceedings of the National Academy of Sciences of the United States of America **92:** 920-924.

ZUPUNSKI, V., F. GUBENSEK and D. KORDIS, 2001 Evolutionary dynamics and evolutionary history in the RTE clade of non-LTR retrotransposons. Molecular Biology and Evolution **18:** 1849-1863.

Vita of John Howard Starnes

Date and place of birth:
April 30th, 1978 in Bowling Green, KY

Educational Institutions attended and degrees already awarded:
M.S., Biology, Western Kentucky University, 2004
B.S., Agriculture Biotechnology, University of Kentucky, 2000

Professional positions held:
Biology Instructor, August 2010-current
Employer: Somerset Community College

Research Assistant, August 2004-July2010
Employer: Plant Pathology Department, University of Kentucky

Graduate Assistant, August 2002- August 2004
Employer: Department of Biology, Western Kentucky University

Substitute Teacher August 2000-May 2002
Employer: Edmonson County Public Schools

Laboratory Assistant, May 2000-August 2000
Employer: Plant Disease Diagnostic Laboratory

Scholastic and professional honors:
KCTCS Merit Bonus Award (KCTCS, 2012)
NSF-ATE Grant to attend Educational Workshop (KCTCS, 2012)
Research Challenge Trust Fund Fellowship, (Univ. of Kentucky, 2004-2011)
Graduate Assistantship, (Western Kentucky University, 2003-2004)
Research Assistantship, (Western Kentucky University, 2002-2003)
Kentucky Academy of Science Botany Grant (Western Kentucky Univ. 2003)
Graduate Student Research Grant (Western Kentucky Univ., 2003)
Howard Hughes Medical Institute Grant (Univ. of Kentucky, 1999)

Professional publications:

J. H. Starnes, D. W. Thornbury, O.S. Novikova, C.J. Rehmeyer, and M.L. Farman. 2012. Telomere-targeted Retrotransposons in the Rice Blast Fungus *Magnaporthe oryzae*: Agents of Telomere Instability. Genetics. 191. 389-406

O. Novikova, J. Starnes, D. Thornbury, M. Farman. 2010. Elucidating mechanisms of telomere instability in Magnaporthe oryzae: retrotransposons running riot. XXVI Fungal Genetics Conference. Fungal Genetics Reports. 58. (Suppl): Abstract 609. Pacific Grove, California.

M. Farman and J. Starnes. 2007. Telomere instability in *Magnaporthe oryzae* caused by highly mobile, telomere-targeted transposons. Fungal Genetics Reports. 54 (Suppl):Abstract 160. XXIV Fungal Genetics Conference. Pacific Grove, California.

J. A. Hewitt, A.J. Meier, J.H. Starnes, P.K. Hamilton, and C.C. Rhoades 2005. Effects of past landuse and initial treatments on *Castanea dentata* seedlings. In, Restoration of American Chestnut to Forest Lands-Proceedings of a Conference and Workshop. Steiner K.C. and Carlson J.E., eds. The North Carolina Arboretum. Natural Resources Report. National Park Service. Washington D.C. pp. 203-210.

Typed Name:
John Howard Starnes