The avirulence (AVR) gene AVR-Pita in *Magnaporthe oryzae* prevents the fungus from infecting rice cultivars containing the resistance gene Pi-ta. A survey of isolates of the *M. grisea* species complex from diverse hosts showed that AVR-Pita is a member of a gene family, which led us to rename it to AVR-Pita1. Avirulence function, distribution, and genomic context of two other members, named AVR-Pita2 and AVR-Pita3, were characterized. AVR-Pita2, but not AVR-Pita3, was functional as an AVR gene corresponding to Pi-ta. The AVR-Pita1 and AVR-Pita2 genes were present in isolates of both *M. oryzae* and *M. grisea*, whereas the AVR-Pita3 gene was present only in isolates of *M. oryzae*. Orthologues of members of the AVR-Pita family could not be found in any fungal species sequenced to date, suggesting that the gene family may be unique to the *M. grisea* species complex. The genomic context of its members was analyzed in eight strains. The AVR-Pita1 and AVR-Pita2 genes in some isolates appeared to be located near telomeres and flanked by diverse repetitive DNA elements, suggesting that frequent deletion or amplification of these genes within the *M. grisea* species complex might have resulted from recombination mediated by repetitive DNA elements.

In certain plant species, different cultivars exhibit differential resistance to individual pathogen isolates through the gene-for-gene surveillance system (Flor 1971), in which a disease resistance (*R*) gene product detects the presence of the pathogen by recognizing a corresponding avirulence (AVR) gene-derived signal molecule. Presence of a matching AVR-R pair triggers rapid local defense responses, which limit the spread of the pathogen to cells immediately surrounding the initial infection site and also may evoke a systemic resistance (Dangl and Jones 2001). Although *R* gene-mediated resistance is highly effective once triggered, such resistance frequently loses its effectiveness in the field when new races capable of evading deployed *R* genes emerge (Kiyosawa 1982; Kolmer 1989; Leach et al. 2001; McDonald and Linde 2002; Mundt 1990, 1991). New pathogen races can emerge through multiple mechanisms, such as modifying the structure or expression of the AVR gene product or removing the AVR genes from their genome. Characterization of the distribution and genome organization of AVR genes within pathogen species will help understand the evolutionary mechanisms underpinning race variation. To this end, in the study reported here, we analyzed an AVR gene called AVR-Pita (Orbach et al. 2000) among isolates from diverse hosts (including both rice pathogens and those that are not pathogenic to rice) in the *Magnaporthe grisea* species complex.

Members of the *M. grisea* species complex (Couch and Kohn 2002; Couch et al. 2005; Hirata et al. 2007), consisting of multiple phylogenetic species, cause diseases in a wide variety of weeds and cultivated cereal crops such as rice, wheat, barley, millet, and maize (Ou 1985). Rice blast, caused by certain *M. oryzae* isolates (Couch et al. 2005), is a classic gene-for-gene system (Silue et al. 1992; Valent and Chumley 1994) in which hundreds of races have been identified (Zeigler et al. 1994). Management of rice blast via breeding resistant cultivars often has had only short-term success due to the frequent appearance of new races that can overcome newly introduced *R* genes (Valent and Chumley 1994; Zeigler et al. 1994). Although the degree of race variation in this fungus remains a controversial subject (Kang et al. 2000), analysis of its AVR genes has begun to provide new insights into potential mechanisms underlying race variation (Bohnert et al. 2004; Couch et al. 2005; Farman et al. 2002; Fudal et al. 2005; Kang et al. 1995, 2001; Orbach et al. 2000; Sweigard et al. 1995; Tosa et al. 2005; Zhou et al. 2007).

The AVR-Pita gene, originally identified in rice isolate O-137, encodes a protein that exhibits similarity to fungal metalloproteases of the deuterolysin family (Orbach et al. 2000). AVR-Pita was predicted to be processed to a 176-amino-acid active form (AVR-Pita$_{176}$) based on the structure of other fungal metalloproteases, and only this processed form functioned to trigger Pi-ta-mediated resistance when expressed directly in rice cells (Jia et al. 2000). AVR-Pita in certain strains appears genetically unstable, frequently producing spontaneous mutants that gain virulence on Pi-ta rice cultivars under laboratory conditions (Valent and Chumley 1991). Molecular analysis of several such mutants identified a variety of mutation types in *AVR-Pita*, including deletion events ranging in size from 100 bp to over 12.5 kb (Orbach et al. 2000), point mutations (Orbach et al. 2000), and the insertion of a transposon (Kang et al. 2001). A survey of *AVR-Pita* among field isolates suggests that similar mechanisms generate virulent strains in the field (Zhou et al. 2007). We report here that *AVR-Pita* is a member of a gene family comprising functional and nonfunctional *AVR* genes. Based on this result, we renamed *AVR-Pita* to *AVR-Pita1*. Potential mechanisms underpinning the evolution of this family are discussed based on genomic contexts of *AVR-Pita1* and two other members in several isolates from diverse hosts.
RESULTS

AVR-Pita is a member of a gene family unique to the M. grisea species complex.

Twenty-nine isolates from diverse hosts (Table 1) were surveyed to determine whether AVR-Pita is present among strains that are not pathogenic to rice (Fig. 1). A phylogenetic tree inferred from sequences of the internal transcribed spacer (ITS) region of ribosomal RNA encoding genes suggests that these 29 isolates belong to at least three distinct lineages (Fig. 2): i) M. oryzae isolates from rice and other grass species, ii) M. grisea isolates from Digitaria spp., and iii) two isolates from Cyperus spp. that most likely belong to a new species distinct from the M. grisea species complex. This phylogeny is consistent with results from previous studies showing that M. oryzae isolates are genetically distinct from the other isolates in this complex (Couch and Kohn 2002; Hirata et al. 2007; Tosa et al. 2005). One isolate from genus Leersia (G-194) and three isolates from genus Pennisetum (G-78, G-123, and G-223) appear distinct from M. oryzae isolates. However, an in-depth phylogenetic analysis with more markers is required to test the hypothesis that these isolates form a lineage distinct from M. oryzae and M. grisea.

A Southern analysis of EcoRI (a restriction site absent in AVR-Pita)-digested genomic DNA of these isolates revealed the presence of zero to three DNA fragments hybridizing to AVR-Pita (Fig. 1A). In addition, there was substantial variation in the degree of hybridization intensities between and within some of the isolates, suggesting the presence of genes with varying degrees of sequence similarity to AVR-Pita. In O-315, an isolate from wild rice, the probe hybridized to three fragments, suggesting that there are multiple copies of AVR-Pita. In contrast, a few M. oryzae isolates from diverse hosts, including rice (O-135), wheat (T-5 and G-158), Eleusine spp. (G-22, G-77, and G-176), Eragrostis spp. (G-17), and Panicum spp. (G-218 and G-219), lacked a strong hybridization signal. G-194 from genus Leersia, one Digitaria isolate (G-11), and both isolates from genus Cyperus (G-229 and G-231) also appeared to lack AVR-Pita. In contrast, several isolates that are distantly related to rice pathogens, such as G-1, G-156, and G-189 from Digitaria spp. and G-78 and G-223 from Pennisetum spp., appeared to carry AVR-Pita. Many of the isolates apparently lacking AVR-Pita produced a weak hybridization signal, suggesting the presence of a gene that is distantly related to AVR-Pita. This survey result suggested that AVR-Pita might be a member of a gene family.

Low-stringency hybridization conditions were used to isolate AVR-Pita homologs from genomic DNA libraries of six field isolates that are nonpathogenic to rice, including two isolates from Digitaria spp. (G-1 and G-213), two from Pennisetum spp. (G-78 and G-223), one from Eragrostis spp. (G-1), and one from Eleusine spp. (G-22). DNA sequence analysis revealed that the AVR-Pita homologs from G-1, G-213, G-78, and G-223 had 92 to 98% DNA sequence identity to AVR-Pita, whereas those from G-17 and G-22 exhibited only 71 to 72% identity to AVR-Pita (Table 2). When the sequenced genome of strain 70-15 (Dean et al. 2005) was searched via BLASTN using the AVR-Pita gene sequence as a query, two different genes were identified. One (locus ID MGG_11081.4 in the gene set version 4; not included in version 5) had 98% DNA sequence identity to AVR-Pita, whereas the other one (locus ID unassigned to date) exhibited 71% DNA sequence identity to AVR-Pita but 97 to 98% identity to the AVR-Pita homologs from G-17 and G-22 (Table 2). The original AVR-Pita and its homologs from G-1, G-78, G-223, and 70-15 (98% identical to AVR-Pita) were renamed to AVR-Pita1. The homolog from G-213 (92% identical to AVR-Pita1) was designated tentatively as AVR-Pita2, and those from G-17, G-22, and 70-15 as AVR-Pita3. Together, these genes form a new gene family designated as the AVR-Pita family.

The EcoRI-digested genomic DNA blot hybridized with AVR-Pita1 was stripped and rehybridized with AVR-Pita2 and AVR-Pita3 (Fig. 1B and C). Because of the high sequence

Table 1. Fungal isolates used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Host</th>
<th>Country of origin</th>
<th>Comments*</th>
</tr>
</thead>
<tbody>
<tr>
<td>O-135</td>
<td>Oryza sativa</td>
<td>China</td>
<td>Valient et al. 1991</td>
</tr>
<tr>
<td>O-137</td>
<td>O. sativa</td>
<td>China</td>
<td>Sweenig et al. 1995</td>
</tr>
<tr>
<td>O-284</td>
<td>O. sativa</td>
<td>Guiana</td>
<td>Guyl, Leung et al. 1988</td>
</tr>
<tr>
<td>G-160</td>
<td>Leersia hexandra</td>
<td>Philippines</td>
<td>Lh-A8401, J. M. Bonman</td>
</tr>
<tr>
<td>O-315</td>
<td>O. longistimnata</td>
<td>Cameroon</td>
<td>CM13, J. L. Notteghem</td>
</tr>
<tr>
<td>T-5</td>
<td>Wheat</td>
<td>Brazil</td>
<td>S. Igarashi (1988)</td>
</tr>
<tr>
<td>G-158</td>
<td>Fera Triticale</td>
<td>Brazil</td>
<td>K. Kneiz (1982)</td>
</tr>
<tr>
<td>G-200</td>
<td>Eleusine indica</td>
<td>Ivory Coast</td>
<td>CD156, J. L. Notteghem</td>
</tr>
<tr>
<td>G-17</td>
<td>Eragrostis curvula</td>
<td>Japan</td>
<td>K76-79, H. Yaegashi (1976)</td>
</tr>
<tr>
<td>G-22</td>
<td>Eleusine coracana</td>
<td>Japan</td>
<td>WGG-FA40, H. Yaegashi (1977)</td>
</tr>
<tr>
<td>G-176</td>
<td>E. coracana</td>
<td>Nepal</td>
<td>NP10-17-4-1-3, H. Kato</td>
</tr>
<tr>
<td>G-219</td>
<td>Panicum maximum</td>
<td>India</td>
<td>IN5, J. L. Notteghem</td>
</tr>
<tr>
<td>G-165</td>
<td>Brachiaria distichya</td>
<td>Philippines</td>
<td>B-D-A8401, J. M. Bonman</td>
</tr>
<tr>
<td>G-218</td>
<td>P. coloratum</td>
<td>Japan</td>
<td>JP31, J. L. Notteghem</td>
</tr>
<tr>
<td>G-220</td>
<td>P. repens</td>
<td>India</td>
<td>IN3, J. L. Notteghem</td>
</tr>
<tr>
<td>G-221</td>
<td>Pennisetum clandestinum</td>
<td>Japan</td>
<td>JP35, J. L. Notteghem</td>
</tr>
<tr>
<td>G-156</td>
<td>Digitaria horizontalis</td>
<td>Brazil</td>
<td>A. S. Prabhu (1986)</td>
</tr>
<tr>
<td>G-11</td>
<td>D. sanguinalis</td>
<td>Japan</td>
<td>A26, H. Yaegashi (1975)</td>
</tr>
<tr>
<td>G-213</td>
<td>D. motusii</td>
<td>Japan</td>
<td>JP34, J. L. Notteghem</td>
</tr>
<tr>
<td>G-123</td>
<td>P. glaucum</td>
<td>United States</td>
<td>84P-19, H. Wells</td>
</tr>
<tr>
<td>G-223</td>
<td>P. typhoides</td>
<td>Burkina Faso</td>
<td>BF17, J. L. Notteghem</td>
</tr>
<tr>
<td>G-219</td>
<td>L. oryzae</td>
<td>Japan</td>
<td>JP36, J. L. Notteghem</td>
</tr>
<tr>
<td>G-229</td>
<td>Cyperus brevifolius</td>
<td>Philippines</td>
<td>PH54, J. L. Notteghem</td>
</tr>
<tr>
<td>G-231</td>
<td>C. rotundus</td>
<td>Philippines</td>
<td>PH52, J. L. Notteghem</td>
</tr>
<tr>
<td>CP987</td>
<td>Laboratory strain</td>
<td>Laboratory strain</td>
<td></td>
</tr>
</tbody>
</table>

* Reference, alternative name, collector, and year of collection.
similarity between AVR-Pita1 and AVR-Pita2, the probes cross-hybridized. However, in some strains (G-220, G-213, G-32, and G-189), AVR-Pita2 hybridized to fragments that had not been detected clearly by AVR-Pita1, suggesting the presence of an additional family member. Unlike AVR-Pita1, AVR-Pita3 was present only in M. oryzae isolates (Fig. 1C, first 17 lanes). Although G-223 (Pennisetum isolate) also hybridized to the probe, given the weak signal it probably is not AVR-Pita3 but a potential new member. The AVR-Pita1 gene consists of four exons (Orbach et al. 2000). AVR-Pita2 and AVR-Pita3 contained the introns at the same positions with consensus splice sites that were identical to those in AVR-Pita1. The AVR-Pita3 gene

### Table 2. Pairwise sequence comparison between AVR-Pita genes and gene products

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AVR-Pita1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G-78</td>
<td>98/92</td>
<td>...</td>
<td></td>
<td></td>
<td>...</td>
<td></td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>G-223</td>
<td>98/94</td>
<td>99/97</td>
<td></td>
<td></td>
<td>99/99</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>G-1</td>
<td>98/94</td>
<td>98/94</td>
<td>99/96</td>
<td></td>
<td>...</td>
<td>98/96</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>70-15</td>
<td>98/94</td>
<td>98/94</td>
<td>99/96</td>
<td></td>
<td>...</td>
<td>98/96</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>AVR-Pita2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G-213</td>
<td>92/84</td>
<td>92/84</td>
<td>91/83</td>
<td>91/83</td>
<td>91/82</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>AVR-Pita3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70-15</td>
<td>71/55</td>
<td>70/57</td>
<td>70/57</td>
<td>70/57</td>
<td>70/56</td>
<td>69/53</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>G-22</td>
<td>72/56</td>
<td>71/57</td>
<td>70/57</td>
<td>70/57</td>
<td>70/56</td>
<td>70/54</td>
<td>97/92</td>
<td>...</td>
</tr>
<tr>
<td>G-17b</td>
<td>71</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td>98</td>
<td>97</td>
</tr>
</tbody>
</table>

* Percent identities of DNA sequences to the left and amino acid sequences to the right of the slash.

* This protein sequence was not used for comparison because of the presence of a premature stop codon in the gene.
in G-17 contained a frameshift mutation near the start codon, resulting in a truncated protein (18 amino acids), if translated. The sequence similarity among members of the gene family at the amino acid level (53 to 99%) is lower than that at the DNA level (69 to 99%) (Table 2).

AVR-Pita1 exhibits sequence similarities to neutral metalloproteases in Aspergillus spp. and Penicillium citrinum that belong to families M35 and M36 in the superfamily of zinc-dependent metalloproteases (Orbach et al., 2000). According to interProScan analysis (Quevillon et al., 2005), all three AVR-Pita gene products contain interPro profile IPR006025 (zinc-binding region signature; corresponding to residues of 173 to 182 in O-137 AVR-Pita1) and belong to the M35 family. To investigate evolutionary relationships between AVR-Pita proteins and fungal members of the M35 family, we searched known fungal proteomes for proteins that are closely related to AVR-Pita proteins (BLASTP E value <1.0 e−3) and belong to the M35 family (discussed below). There was substantial variation in the number of putative M35 proteins in different fungi (e.g., one in Fusarium graminearum versus seven in Coccidioides immitis). In the 70-15 genome, three putative M35 proteins were identified. Similar to AVR-Pita gene products, two of them (MGG_03029.5 and MGG_10927.5) contain a predicted signal peptide at their N-terminus and interPro Profile IPR006025.

A phylogenetic analysis of putative M35 proteins revealed three clusters (bootstrap support >70%) that contain orthologs from at least three species (Fig. 4). MGG_10927.5 was used as a probe to investigate its distribution in the M. grisea species complex (Fig. 1D). Unlike members of the AVR-Pita family, this gene seems to be present in all isolates, with the exception only of G-229. Putative orthologs of MGG_10927.5 were found in Neurospora crassa (NCU05071.2) and F. graminearum (FG08289.1). However, no fungi sequenced to date seem to carry orthologs of members of the AVR-Pita family, suggesting that the gene family may be unique to the M. grisea species complex.

**Functional analysis of AVR-Pita family members.**

Clones carrying individual family members were introduced via transformation into CP987, a laboratory strain lacking AVR-Pita1, to determine whether they confer avirulence. The AVR-Pita1 gene from O-137, G-1, G-78, and G-223 and the AVR-Pita2 gene from G-213 converted CP987 to avirulence toward rice cultivars carrying Pi-ta, but the AVR-Pita3 gene from G-17 and G-22 failed to do so (data not shown). Possible reasons for the failure of the G-22 AVR-Pita3 gene to confer avirulence would be improper expression of the gene or production of a protein that may not interact with Pi-ta. To investigate the underlying cause of the G-22 AVR-Pita3’s failure to confer avirulence, we produced a series of chimeric genes. First, to test whether the G-22 AVR-Pita3 coding sequence (designated as C\_AVR-Pita3) can produce a functional avirulence protein, we fused it to the O-137 AVR-Pita1 and G-213 AVR-Pita2 gene promoters (designated as P\_AVR-Pita1 and P\_AVR-Pita2, respectively), producing P\_AVR-Pita1::C\_AVR-Pita3.

---

**Table 3.** Sequence comparison among members of the AVR-Pita family. The number (1, 2, or 3) next to the strain name indicates AVR-Pita1, AVR-Pita2, or AVR-Pita3, respectively. AVR-Pita1 from O-137, G-1, G-78, and G-223 and AVR-Pita2 from G-213 confer avirulence (AVR), whereas AVR-Pita3 from 70-15, labeled as 70-15(3), and G-22(3), lack avirulence function (avr). Although the avirulence function of AVR-Pita1 from 70-15, labeled as 70-15(1), was not determined (n.d.) in this study, considering that 70-15 is virulent on Pi-ta-carrying cultivars, the gene probably is nonfunctional as an AVR gene. Sequence positions were numbered according to O-137 AVR-Pita1. Asterisks indicate the protease motif that is shared by the majority of zinc-dependent metalloprotease (interPro profile IPR006025; corresponding to residues of 173 to 182 in O-137 AVR-Pita1). [GSLTVV]-[x2]-H-E-[LIVMFYW]-[DEHRKP]-H-x-[LIVMFYLGSG], in which [] indicates that any one of the listed amino acid residues is allowed, x indicates that any residue is allowed, and {} indicates that any one of the listed residues is not allowed. Boldfaced residues above (two histidines and one glutamic acid) correspond to putative zinc ligands and the active site residue, respectively. The putative signal peptide sequence is overlaid with an arrow marking its predicted cleavage site.
and P\textsubscript{avr-Pita3}$\cdot$C\textsubscript{avr-Pita3}. Functionality of the G-22 AVR-Pita3 gene promoter also was tested by fusing it to the O-137 AVR-Pita1 and G-213 AVR-Pita2 coding sequences (yielding P\textsubscript{avr-Pita1}$\cdot$C\textsubscript{avr-Pita1} and P\textsubscript{avr-Pita2}$\cdot$C\textsubscript{avr-Pita2}, respectively). Five additional constructs, including P\textsubscript{avr-Pita3}$\cdot$C\textsubscript{avr-Pita3}, P\textsubscript{avr-Pita3}$\cdot$C\textsubscript{avr-Pita2} and P\textsubscript{avr-Pita2}$\cdot$C\textsubscript{avr-Pita2}, were produced as controls to determine whether the fusion process affected gene function.

Individual chimeric constructs were introduced to CP987, and randomly chosen transformants were analyzed by polymerase chain reaction (PCR) to confirm that the integrated genes were intact. Five to six confirmed transformants with each construct were inoculated on Pi-ta rice, and host responses were scored. The P\textsubscript{avr-Pita1}$\cdot$C\textsubscript{avr-Pita1} and P\textsubscript{avr-Pita2}$\cdot$C\textsubscript{avr-Pita2} constructs successfully conferred avirulence, as did P\textsubscript{avr-Pita3}$\cdot$C\textsubscript{avr-Pita2} and P\textsubscript{avr-Pita3}$\cdot$C\textsubscript{avr-Pita1}. As expected, the P\textsubscript{avr-Pita3}$\cdot$C\textsubscript{avr-Pita1} construct failed to confer avirulence. These results showed that individual modules retained their original functionality after the construction process. All transformants carrying either P\textsubscript{avr-Pita1}$\cdot$C\textsubscript{avr-Pita1} or P\textsubscript{avr-Pita2}$\cdot$C\textsubscript{avr-Pita2} caused typical blast symptoms on Pi-ta rice, suggesting that G-22 AVR-Pita3 encodes a protein that does not trigger Pi-ta-mediated resistance. Interestingly, two of six transformants carrying the P\textsubscript{avr-Pita3}$\cdot$C\textsubscript{avr-Pita3} construct and four of six transformants carrying P\textsubscript{avr-Pita2}$\cdot$C\textsubscript{avr-Pita2} construct caused intermediate resistance responses, suggesting the expression of AVR-Pita1 and AVR-Pita2. However, the remaining transformants caused blast disease on Pi-ta rice. Considering that all of the transformants had an intact transgene, differences in gene expression among transformants with P\textsubscript{avr-Pita3}$\cdot$C\textsubscript{avr-Pita3} or P\textsubscript{avr-Pita2}$\cdot$C\textsubscript{avr-Pita2} probably were caused by position effects (i.e., genome sequences flanking individual insertion sites differentially affecting expression of the transgene).

Some members of the AVR-Pita family are telomeric.

Similar to the O-137 AVR-Pita1 gene, which is located near a telomere (Orbach et al. 2000), the AVR-Pita1 gene in D. taria isolate G-1 also appears to be telomeric based on chromosome walking via inverse PCR. However, its orientation relative to the telomeric repeat was reversed (Fig. 5A); in O-137 AVR-Pita1, the telomeric repeat is located 48 bp downstream from its stop codon, whereas the telomeric repeat (TTAGGG\textsuperscript{20}) is located at the 5’ region of G-1 AVR-Pita1 (1,689 bp upstream from the start codon). Considering the telomeric location of the AVR-Pita1 gene in these two distantly related isolates (Fig. 2), we hypothesized that members of the gene family in other isolates also might be telomeric. The close linkage between the telomere and G-223 AVR-Pita1 and G-213 AVR-Pita2 was confirmed by cloning of fosmid genomic clones containing these genes. Restriction analysis of these clones indicated that, in both G-213 and G-223, the distance between the gene and the telomeric repeat appeared to be much greater than that of the AVR-Pita1 gene in O-137 and G-1, but less than 30 kb (data not shown). No clones containing AVR-Pita3 hybridized to the telomeric repeat.

We used genetic and physical maps available in the genome database of 70-15 to find the chromosomal positions of AVR-Pita1 and AVR-Pita3 in this strain. The AVR-Pita1 gene was found in the contig 2.2332 (3,251 bp) that belongs to the supercontig 124 (5,281 bp). Based on version 4 of the genome assembly, supercontig 124 has not yet been mapped on any linkage group. The AVR-Pita3 gene was found in the supercontig 179 (version 5), mapped on chromosome 7. Recently, a new chromosome 7 sequence assembly was released (approximately 4 mb; GenBank accession number CM000230). The AVR-Pita3 gene is separated from the telomeric repeat by more than 300 kb, indicating that it is not telomeric, and is located within one of the transposable element clusters (Thon et al. 2006).

Multiple transposable elements are closely associated with the AVR-Pita family.

Members of the AVR-Pita family are closely associated with diverse transposons (Fig. 5). They include i) two different inverted terminal repeat (ITR) transposons that exhibit high similarity to Pot2 (Kachroo et al. 1994) and Pot3 (Farman et al. 1996), respectively; ii) a highly degenerated ITR transposon (tentatively termed ψ transposon); iii) a solo long-terminal repeat (LTR) derived from retrotransposon Pyret (Nakayashiki et al. 2001); iv) two novel retrotransposons; and v) three undefined, short repetitive DNA elements, including MGR619 (79 bp), MGR608 (126 bp), and REPI (188 bp). The presence of MGR619, MGR608, and REPI was intriguing because they are closely associated with other AVR genes, including PWL (Kang et al. 1995) and AVR1-5039 (Farman et al. 2002). The MGR619 and MGR608 elements are highly polymorphic and dispersed sequences (Hamer et al. 1989) and frequently appear together as a contiguous sequence (Kang et al. 1995; Orbach et al. 2000).
Some of the MGR619/MGR608 and REP1 elements flanking members of the AVR-Pita gene family were attached to sequences exhibiting significant homology to reverse transcriptase of LTR-type retrotransposons. In addition, the lengths of MGR619/MGR608 (205 bp) and REP1 (188 bp) were similar to those of fungal retrotransposon LTRs (ranging from 198 bp to 596 bp in size) (Galun 2003). To determine whether the MGR619/MGR608 and REP1 elements were solo LTRs (i.e., LTRs missing the internal domains of retrotransposon) derived from novel retrotransposons, we searched the genome database of 70-15 using MGR619/MGR608 and REP1 as queries. In many positive contigs, the internal domains of LTR retrotransposon (e.g., the gag and pol genes) were sandwiched between copies of MGR619/MGR608 or REP1. However, none of the positive contigs contained an intact open reading frame of the pol gene, suggesting that they correspond to degenerate retrotransposons.

The borders of the LTRs corresponding to MGR619/MGR608 and REP1 were defined based on sequence alignment among multiple copies of retrotransposons from different contigs (data not shown). We used one REP1-associated retrotransposon and a retrotransposon copy associated with MGR619/MGR608 to determine whether they correspond to any LTR-retrotransposons previously characterized (Dean et al. 2005; Thon et al. 2006). The REP1-associated copy exhibited 91.8% sequence identity with RETRO6-1 and the MGR619/MGR608-associated copy had 91.7% sequence identity with RETRO7-1. The REP1 element (158 bp) was identical in sequence to that of the RETRO6-1 LTRs (193 and 194 bp, respectively, with one additional nucleotide in 3' LTR), suggesting that REP1 is a truncated derivative from RETRO6-1. Similarly, the MGR619/608 element (205 bp) is a deletion derivative of the RETRO7-1 LTRs (215 bp).

**Comparison of the regions flanking members of the AVR-Pita family.**

Sequence similarity at the AVR-Pita1 locus among different strains, and between AVR-Pita1 and AVR-Pita2, extended beyond the coding sequence at both the 5' and 3' sides but was interrupted frequently by retrotransposons or solo LTRs (Figs. 5A and 6). For instance, between G-213 AVR-Pita2 and O-137 AVR-Pita1, the synteny in the promoter region was interrupted by insertion of a 192-bp solo LTR of RETRO6-1 in the AVR-Pita2 promoter at the position corresponding to –332 of the O-137 AVR-Pita1 promoter. The synteny continued after this solo LTR but was interrupted again, at position –447 in the O-137

---

**Fig. 5.** Comparison of the genomic context among members of the AVR-Pita family in seven field isolates and one laboratory strain. GenBank accession numbers for the loci used here are as follows: O-137 (AF2207841), G-1 (DQ855953), G-223 (DG855954), G-78 (DG855955), G-213 (DG855956), G-17 (DG855957), and G-22 (DG855958). For 70-15 sequences, contig 2.2232 and GenBank accession number CM000230 (sequences between 3668019 and 3676012) were used. A putative coding sequence is indicated by an open box with a gene name. A highly degenerate transposon is indicated by ψ. The box with T indicates the telomeric repeat (not drawn in scale). Long-terminal repeat (LTR) retrotransposons RETRO6-1 and RETRO7-1 and solo LTRs derived from them are marked with boxes filled with different patterns. A solo LTR (or its derivative) adjacent to another LTR is indicated in a staggered manner. A 5-bp target site duplication (G-213) is marked with a pair of filled triangles. The orientation of transcription for individual genes is indicated by an arrow. Gray areas between genes indicate the syntenic regions with >96% sequence identity. **A**, The AVR-Pita1 and AVR-Pita2 loci among six isolates were compared. **B**, The AVR-Pita3 locus was compared among three different isolates.
AVR-Pita1 promoter, by the presence of a solo LTR of RETRO7-1 (Fig. 6). At the corresponding site of AVR-Pita2, a partial solo LTR of RETRO7-1 was found (oriented opposite to that in O-137), which was followed immediately by a copy of RETRO7-1. At the 3′ end, O-137 AVR-Pita1 contained the telomeric repeat 48 bp downstream from the stop codon, whereas the G-213 AVR-Pita2 was flanked by a copy of RETRO7-1 that exists in a head-to-head orientation with the RETRO7-1 element at the 5′ side of the gene (98% sequence identity between them).

The AVR-Pita1 gene in other strains also was surrounded by retrotransposons and solo LTRs. A truncated copy of RETRO6-1 was located between the telomeric repeat and the 5′ end of the G-1 AVR-Pita1 gene (Fig. 5A); when compared with an intact copy of RETRO6-1 (5,871 bp), it contained only the 3′ 1,181 bp of RETRO6-1, immediately followed by the telomeric repeat at the 5′ end. Although AVR-Pita1 in other strains and AVR-Pita2 in G-213 do not have RETRO6-1 at their 5′ region, the presence of a solo RETRO6-1 LTR suggests that RETRO6-1 once existed at this location but has been deleted. Numerous stop codons exist in the internal domains of the AVR-Pita-associated RETRO6-1 and RETRO7-1 elements, indicating that they have degenerated.

Between AVR-Pita1/AVR-Pita2 and AVR-Pita3, noticeable sequence identity (65%) existed only up to approximately 180 bp upstream from the translation start codon but no significant similarity was found downstream of the stop codon. However, at the AVR-Pita3 locus in G-17, G-22, and 70-15, there was strong sequence conservation (more than 93% identity) at both sides of the gene (Fig. 5B). The synteny between G-22 and 70-15 at the AVR-Pita3 locus was disrupted by transposons (G-22 contained a highly degenerated transposon whereas 70-15 carried Pot3 and a solo-LTR of Pyret) as well as deletions (or insertions) of two DNA segments; the AVR-Pita3 locus in G-17 and 70-15 lacks a 197-bp segment at the 5′ region of the AVR-Pita3 locus in G-22. A 329-bp segment at the 5′ region of the gene in 70-15 appears to be missing in both G-17 and G-22. BLAST searches with these segments as queries did not reveal any significant matches in GenBank, nor did these segments exhibit any distinct structural features such as direct repeats at their ends or flanking regions.

Signs suggesting multiple transpositions and subsequent rearrangements near members of the gene family were abundant. A partial solo LTR of RETRO7-1 immediately flanked the 3′ end of the truncated RETRO6-1 in G-1 and a solo LTR of RETRO6-1 in G-223 and 70-15 (Fig. 5A). In contrast, at the same location in O-137, only an intact solo LTR of RETRO7-1 was present (Fig. 5A). These data suggest that an LTR of RETRO7-1 existed at this location in the strain ancestral to G-1, G-223, and O-137, and transposition of RETRO6-1 into this LTR occurred later in the lineage leading to G-1 and G-223 but not in the lineage leading to O-137. This RETRO6-1 element in G-1 and G-223 subsequently has undergone additional changes (e.g., chromosomal breakage and addition of the telomeric repeat in G-1 and deletion of the internal domain of RETRO6-1 probably mediated by recombination between its LTRs in G-223). The orientation of RETRO7-1 in the 3′ region of the AVR-Pita1 gene in G-1 and G-223 was opposite to that in G-78 and G-213, suggesting independent transposition events at this location.

**DISCUSSION**

Similar to the PWL AVR gene family in the *M. grisea* species complex (Kang et al. 1995), members of the AVR-Pita family are widely distributed among isolates from diverse hosts (Figs. 1 and 2), including those that are not pathogenic to rice. Based on the genomic context and distribution of members of the AVR-Pita family, we discuss their function, evolution, and variation.

**Do members of the AVR-Pita gene family have more than AVR function?**

An intriguing question about AVR genes in general is why, during evolution, plant pathogens have maintained genes that limit their ability to infect potential hosts. Accumulating evidence indicates that, in both bacterial and fungal pathogens, the maintenance of certain AVR genes is due to their function as a virulence factor in hosts lacking corresponding R genes. A number of AVR genes in bacterial pathogens have been shown to contribute to virulence (Abramovitch et al. 2003; Axtell et al. 2003; Bai et al. 2000; Chang et al. 2000; Chen et al. 2004; Greenberg and Vinatzer 2003; Leach et al. 2001; Lorang et al. 1994; Marois et al. 2002; Ritter and Dangl 1995; Yang et al. 1994, 1996). Fungal AVR genes demonstrating a dual role include *ecp2* (Laugé et al. 1997) and *Avr4* (van den Burg et al. 2006; van Esse et al. 2007) in *Cladosporium fulvum*, *nip1* in *Rhynchosporium secalis* (Knogge 1996; Weversiep et al. 1991),

![Image](image_url)

**Fig. 6.** Comparison of the promoter and coding regions of O-137 AVR-Pita1 (-489 to +82) and G-213 AVR-Pita2 (-650 to +884). Numbers above the boxes indicate the positions defined relative to the translation start site (A of ATG as +1). The ratios indicate the number of avirulent transformants among the total number of transformants containing the corresponding subclones in the complementation analyses (derived from Orbach et al. 2000). Percent identities between the corresponding regions are indicated. A partial copy of RETRO7-1 long-terminal repeat (LTR) (43 bp) and a complete copy of RETRO6-1 LTR (192 bp), present at the 5′ ends of the O-137 AVR-Pita1 promoter and the G-213 AVR-Pita2 promoter, respectively, are indicated. A putative 6-bp target site duplication (ATTACT and ATTACC with a change at the sixth position) generated during transposition is highlighted.
and found variants that fall between AVR-Pita2 sequenced generated after the separation of gene. However, the degree of sequence variation within possibility that they are highly divergent alleles of the same lidity of this designation has not yet been firmly proven. We were compared), and a similar pattern existed within genes from strains distantly related to rice pathogens. Interestingly, the in the amplification of the gene family member. This mechanism appears to have involved a mechanism for gene duplication. Gene conversion initiated ruling out retrotransposition (i.e., insertion of a reverse tran-

Possible mechanisms underlying the evolution of the AVR-Pita family. The presence of both AVR-Pita1 and AVR-Pita2 in several isolates (Fig. 2) suggests that they were derived from an ancestral gene duplication event. Although we designated AVR-Pita1 and AVR-Pita2 as distinct members of the gene family, the va-

population of the M. grisea species complex but has been com-

ments through recombination between dispersed repeats located at different chromosome locations is a possible mechanism for the generation of a new gene family member. This mechanism appears to have involved in the amplification of the PWL2 AVR gene (S. Kang, unpul-

in haustoria, suggesting their roles in virulence (Catanzariti et al. 2006). The Avr2 gene product of C. fulvum may contribute to virulence by inhibiting apoplastic cysteine proteases in to-

neighboring sequences at the ends of telomeres. Most subtelomeric regions in the M. oryzae isolates were identified in a weeping lovegrass pathogen (Valent et al. 1991). Gene-for-gene interactions of rice culti-

bers suggest such roles. Rice pathogens that entirely lack members of this family were rarely found among approximately 200 isolates from sev-

eral countries (Fig. 1; unpublished data), suggesting the presence of positive selection pressure for their maintenance. Ex-

pression from AVR-Pita1 was not detected under various axenic culture conditions but was induced in rice during an early stage of infection (B. Valent and S. Kang, unpublished data). Al-

though new races that specifically overcame resistance con-

ferred by Pi-ta in rice cv. Katy have been isolated from rice fields in the United States, this cultivar has been stably resis-
tant to rice blast since its introduction in the late 1980s (Jia et al. 2004). Preservation of functional members of the gene family in M. grisea isolates that are not pathogenic to rice, such as G-1, G-78, G-213, and G-223, also suggests that this family may carry out functions unrelated to avirulence factors in such strains. Presence of rice-specific AVR genes among isolates nonpathogenic to rice is not unique to the AVR-Pita family (Tosa et al. 2005; Valent et al. 1991; Yaegashi and Asaga 1981). For instance, AVR genes AVR1-CO39, AVR1-M201, and AVR1-YAMO were identified in a weeping lovegrass pathogen (Valent et al. 1991). Gene-for-gene interactions of rice culti-
vars with M. oryzae likely have evolved through the specific recognition by rice of preexisting fungal gene products that originally had roles other than triggering rice defense as aviru-

lence factors.

in rice cv. Katy have been isolated from rice

pleteness deleted in M. grisea. The latter possibility (requiring multiple independent deletions of the gene in different host-
specific lineages) seems unlikely. The AVR1-CO39 gene also appears to be absent in M. grisea (Tosa et al. 2005). Interest-
ingly, AVR1-CO39 is absent among rice isolates (Farman et al. 2002).

The AVR-Pita1 and AVR-Pita2 genes in certain strains seem to be located in subtelomeric regions, at distances ranging from 48 bp up to approximately 30 kb from the telomeric repeat. Variation in the distance and orientation between the telomeric repeat and members of the gene family suggests the dynamic nature of the subtelomeric regions associated with AVR-Pita1 and AVR-Pita2. Other subtelomeric regions in M. oryzae also appear highly dynamic, as illustrated by the telomere-linked helicase (TLH) gene family (Gao et al. 2002; Rehmeyer et al. 2006). Most TLH family members are located within 10 kb of the telomeric repeat, and undergo frequent deletion and ampli-

fication events, which likely have been mediated by inter-

chromosomal recombination between the repetitive elements flanking them. It remains to be determined whether similar recombination mechanisms contribute to the proliferation of the AVR-Pita family. Telomeres of the M. oryzae isolates from turfgrass also exhibited hypervariability (Farman and Kim 2005). Frequent deletions and rearrangements at subtelomeric regions have been observed in diverse eukaryotes (Carlson et al. 1985; Charron and Michels 1988; Heather and Trask 2002). In general, those genes whose frequent variation can confer an advantage for adapting in certain niches often are present in telomeric regions. In Saccharomyces cerevisiae, members of the SUC and MAL gene families that are involved in metabolizing specific carbon sources (sucrose and maltose, respectively) were found at subtelomeric regions (Carlson et al. 1985; Charron and Michels 1988). The presence and locations of these subtelomeric genes vary widely from strain to strain, with individual strains carrying only a subset of gene family members at different chromosomes, suggesting frequent re-

combination between subtelomeric regions. Similarly, subte-

lomeric regions of the human malarial parasite Plasmodium falciparum undergo frequent ectopic recombination, promot-
ing the variation of var genes in these regions and thus helping the parasite evade the host immune system (Freitas-Junior et al. 2000).

In addition to the AVR-Pita family, M. oryzae has several AVR genes, such as AVR1-Ka86, AVR1-MedNoi, and PWL1, that are mapped to the telomere (Dioh et al. 2000; Kang et al. 1995). Given the dynamic nature at the subtelomeric regions, the presence of AVR genes in these regions may provide an advantage to the fungus by helping it evade newly deployed R genes via frequent loss or modification of AVR. Considering the possibility that certain AVR genes also might function as a virulence factor, we speculate that two opposing selection pressures (increased virulence on host cultivars lacking the corresponding R gene versus incompatibility with those carrying the R gene) have driven the direction of change (amplifica-

tion versus deletion or modification) in fungal populations. We should note that certain AVR genes in M. oryzae, such as AVR1-Mara (Mandel et al. 1997), AVR1-CO39 (Farman and Leong 1998), PWL2 (Sweigard et al. 1995), AVR-Pik (Yasuda et al. 2005), and ACE1 (Bohnert et al. 2004), are not telomeric.

Replicative DNA elements have been implicated as a major source for genetic variation, which can modify the host range of affected strains. A transposition event caused a loss-of-function mutation in two AVR genes of M. oryzae, leading to a change in virulence spectrum (Fudal et al. 2005; Kang et al. 2001; Zhou et al. 2007). Repetitive elements also can cause genome rearrange-

ments through recombination between dispersed repeats located
on the same or different chromosomes, resulting in deletion, duplication, inversion, and translocation, depending on the relative orientation and position of recombining repeats. Some of these changes potentially can affect host range by mutating an AVR gene at or near the recombination site. Members of the PWL AVR gene family are flanked by diverse transposons, and such associations appear to be responsible for deletions or amplifications in the gene family (Kang et al. 1995, 2000; Sweigard et al. 1995). The AVR1-CO39 gene also is closely associated with transposons that likely were responsible for the deletion of the gene among rice-pathogenic isolates of M. oryzae (Farman et al. 2002; Tosa et al. 2005). Two AVR genes in Leptosphaeria maculans, AavrLm1 and AavrLm6, are embedded in long chromosomal regions mainly consisting of several types of LTR retrotransposons and their remnants (Fudal et al. 2007; Gout et al. 2006).

Transposons also can modify the expression pattern of AVR genes near their insertion sites. An insertion of a Pot3 ITR transposon into the promoter region of AVR-Pita1 (–302 bp) abolished its avirulence activity (Kang et al. 2001). Since the discovery that LTRs carry promoter and enhancer motifs (Sverdlov 1999), many studies have shown that LTRs can influence the expression of adjacent genes both spatially and temporally (Brosius 1999). The AVR-Pita1 and AVR-Pita2 genes have one or more LTRs at their 5′ region; the O-137 AVR-Pita1 gene has three solo LTRs in its 5′ region (Fig. 5). A deletion analysis of the putative promoter region of the O-137 AVR-Pita1 gene (Orbach et al. 2000) showed that the two tandem copies of RETRO6-1 solo LTR were not essential for avirulence activity of the gene. A small part of a RETRO7-1 solo LTR (5′ 30 bp) is included in a region that is required for avirulence activity (Fig. 6, between –476 and –388), but its absence in the AVR-Pita2 and G-78 AVR-Pita2 promoter regions (Fig. 5) suggests that this sequence might not be essential. A solo LTR of RETRO6-1 (192 bp) was inserted into an AVR-Pita2 promoter region (at –332 in the O-137 AVR-Pita1 promoter) that is essential for the activity of O-137 AVR-Pita1. Because measuring avirulence activity based on plant infection is not a sensitive assay, determining if and how this solo LTR affects the expression of AVR-Pita2 requires gene expression analysis.

### Table 3. Polymerase chain reaction (PCR) primers used in this study

<table>
<thead>
<tr>
<th>Target</th>
<th>Primer</th>
<th>Sequence (5′→3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal transcribed spacer (ITS)</td>
<td>ITS1</td>
<td>TCCGTAAGTTGACACCGCGG</td>
</tr>
<tr>
<td></td>
<td>ITS4</td>
<td>TCCTCGCTTTATGGATAGC</td>
</tr>
<tr>
<td>Insert of pSK1059</td>
<td>LF9</td>
<td>GCGATTTGCCCCTTCACC</td>
</tr>
<tr>
<td></td>
<td>Pita1</td>
<td>CCTCTATGTGATTGTTGA</td>
</tr>
<tr>
<td>Insert of pSK1183</td>
<td>Fun1</td>
<td>ATAATAATCGA.CGCGTTCGCC</td>
</tr>
<tr>
<td></td>
<td>Fun2-2</td>
<td>GCCGTCGAATCTTCCCTTTAT</td>
</tr>
<tr>
<td>Inverse PCR</td>
<td>ip-1</td>
<td>TCTGTCCTCCGACCTTTGA</td>
</tr>
<tr>
<td></td>
<td>ip-2</td>
<td>GCCGCGTTAAAGGTAATAT</td>
</tr>
<tr>
<td>Insert of pSK2544</td>
<td>MG10927-F</td>
<td>ATCGAATTCTGCTATGACAAAGCTGG</td>
</tr>
<tr>
<td></td>
<td>MG10927-R</td>
<td>ACCGGTCGACCACTTTCATAG</td>
</tr>
<tr>
<td>PAVR-Pita1</td>
<td>Fs-P1</td>
<td>CGGAATTCGCGGAGTGTTCTG</td>
</tr>
<tr>
<td></td>
<td>Fs-P2</td>
<td>CGGAATTCGGTTACGAGATTC</td>
</tr>
<tr>
<td>CAVR-Pita1</td>
<td>Fs-C1</td>
<td>CGGAATTCGCGGAGTGTTCTG</td>
</tr>
<tr>
<td></td>
<td>Fs-C2</td>
<td>CGGAATTCGCGGAGTGTTCTG</td>
</tr>
<tr>
<td>PAVR-Pita2</td>
<td>Fw-P1</td>
<td>CGGAATTCGCGGAGTGTTCTG</td>
</tr>
<tr>
<td></td>
<td>Fw-P2</td>
<td>CGGAATTCGCGGAGTGTTCTG</td>
</tr>
<tr>
<td>CAVR-Pita2</td>
<td>Fs-C1</td>
<td>Same as the Fs-C1</td>
</tr>
<tr>
<td></td>
<td>Fs-C2</td>
<td>Same as the Fs-C1</td>
</tr>
<tr>
<td>PAVR-Pita1</td>
<td>nF-P1</td>
<td>CGGAATTCGCGGAGTGTTCTG</td>
</tr>
<tr>
<td></td>
<td>nF-P2</td>
<td>CGGAATTCGCGGAGTGTTCTG</td>
</tr>
<tr>
<td>CAVR-Pita1</td>
<td>nF-C1</td>
<td>CGGAATTCGCGGAGTGTTCTG</td>
</tr>
<tr>
<td></td>
<td>nF-C2</td>
<td>CGGAATTCGCGGAGTGTTCTG</td>
</tr>
</tbody>
</table>

* Underlined sequences correspond to restriction enzyme sites introduced for cloning purpose: EcoRI (GAATTCC), XbaI (TCTAGA), BamHI (GATCTC), and SphI (GCATGC).

### MATERIALS AND METHODS

#### Strains and growth conditions.

Host and geographical origins of the 29 field isolates used in this study are described in Table 1. The laboratory strain 70-15 was derived from a cross between rice-pathogenic isolate Guy 11 and a weeping lovegrass (Eragrostis curvula)-pathogenic isolate, followed by backcrosses to Guy 11 (Chao and Ellingboe 1991; Leung et al. 1988). Another laboratory strain, CP987, is a spontaneous mutant that lacks functional AVR-Pita1 and, thus, is virulent on Pi-ta rice (Orbach et al. 2000). *Escherichia coli* XL1-blue MRF was used for maintaining plasmids. *Agrobacterium tumefaciens* EHA105 (Klee 2000) was used for fungal transformation. Fungal cultures were grown at 25°C on oatmeal agar plates to produce conidia for infection assays. Mycelial cultures for genomic DNA extraction were grown in liquid complete medium (6 g of yeast extract, 6 g of casein acid hydrolysate, and 10 g of sucrose per liter).

#### Phylogenetic analyses.

The ITS region was amplified from each strain listed in Table 1 with the primers ITS1 and ITS4 (Table 3) using the FailSafe PCR system (Epicentre Technologies, Madison, WI, U.S.A.). The PCR cycling program included an initial denaturation for 2 min at 94°C, 30 cycles of 1 min of denaturation at 94°C, 1 min of annealing at 55°C, and 1 min of extension at 72°C; followed by a final extension for 10 min at 72°C. PCR products were purified using QIAquick spin columns (Qiagen, Valencia, CA, U.S.A.) and sequenced on both strands with the same primers used in the PCR amplification. ITS sequences were aligned using the Clustal W program (Thompson et al. 1994) with default parameters. A phylogenetic tree was constructed using the neighbor-joining method (Saitou and Nei 1987) in the MEGA3 program (version 3.1) (Kumar et al. 2004) with the following parameters: complete deletion of gaps, Jukes-Cantor model, both transitions and transversions substitutions included, and 1,000 bootstrap replicates.

Proteins related to AVR-Pita in other fungi were identified through the following steps. First, multifungi BLASTP (E value <1 e-05) against a database of “all fungal proteins” at the
Broad Institute using O-137 AVR-Pita1, G-213 AVR-Pita2, and G-22 AVR-Pita3 as queries. This led to the identification of 23 proteins, including four from Aspergillus nidulans, two from Botrytis cinerea, seven from Coccidioides immitis, one from F. graminearum, five from M. grisea (not including AVR-Pita1 and AVR-Pita3), two from N. crassa, and two from Sclerotinia sclerotiorum. These protein sequences were analyzed using InterProScan to detect the presence of signatures present in AVR-Pita proteins (IPR001384, IPR006025, and signal peptide). Two M. grisea proteins (MGG_03808.5 and MGG_12056.5) lacked IPR001384 and IPR006025 profiles and were excluded from further analyses. From the MEROPS database, a manually curated information resource for proteolytic enzymes (Rawlings et al. 2006), four representative M. grisea members (or called type peptidase) were retrieved; these include penicillopsin (M35.001; Penicillium citrinum), deuterolysin (M35.002; A. fumigatus), extracellular peptidase (M35.003; Aeromonas hydrophila), and peptidyl-Lys metallo-peptidase (M35.004; Armillaria mellea). After aligning protein sequences, they were adjusted manually based on conserved sites of the type peptidases. Because of significant sequence divergence at the N- and C-terminal regions of these proteins, only parts that overlap with the peptidase units in the type peptidases were included. A phylogenetic tree was constructed using the neighbor-joining method in the MEGA3 program (version 3.1) with the following parameters: complete deletion of gaps, Poisson correction model, and 1,000 bootstrap replicates.

**Southern analysis.**

Fungal genomic DNA was prepared and purified by CsCl gradient centrifugation as previously described (Kang 2001). EcoRI-digested genomic DNA was separated by 0.8% agarose gel electrophoresis and blotted to Hybond N+ membrane (Amerham, Piscataway, NJ, U.S.A.) according to the manufacturer’s instructions. The blot was hybridized successively with 32P-labeled probes of AVR-Pita1 (0.5-kb SmI fragment of pSK1059), AVR-Pita2 (1.0-kb EcoRI fragment of pSK1183), AVR-Pita3 (0.6-kb KpnI-BglII fragment of pSK924), and MGG_10927.5 (1.3-kb EcoRI-XbaI fragment of pSK2544) prepared by random priming; each probe DNA fragment was precipitated by adding 10 μl of 3 M sodium acetate, pH 5.5, and 200 μl of 100% ethanol and incubating for 30 min at –20°C; pelleted by centrifugation; washed with 70% ethanol; air dried; and resuspended in 50 μl of Tris-EDTA (TE) buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). The digested DNA was self-ligated for 16 h at 16°C in a 500-μl reaction, ethanol precipitated, and subsequently resuspended in 30 μl of TE buffer. PCR was performed in a 20-μl reaction mixture containing 10 μl of template DNA and 2 μl (10 pmol/μl) of primers IP-1 and IP-2 (Table 3), using the FailSafe PCR system. The PCR cycling program consisted of an initial denaturation for 5 min at 95°C, 30 cycles of 30 s of denaturation at 95°C, 30 s of annealing at 60°C, and 2 min of extension at 72°C, followed by a final extension for 10 min at 72°C. PCR products were cloned into pGEM-T Easy vector (Promega Corp.).

Plasmid clones were sequenced using primers based on the cloning vector or the EZ::TN transposon insertion system.

**Table 4.** Key plasmids used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Description, reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSK646</td>
<td>pCB780; 6.5-kb BglII blunt telomere fragment from 4224-7-8 in BanHI-EcoRV sites of pBluescript SK+ (Orbach et al. 2000)</td>
</tr>
<tr>
<td>pSK647</td>
<td>3.0-kb EcoRI fragment carrying AVR-Pita1 from G-1 in EcoRI site of pGEM3Zf</td>
</tr>
<tr>
<td>pSK648</td>
<td>5.0-kb EcoRI fragment carrying AVR-Pita1 from G-78 in EcoRI site of pGEM3Zf</td>
</tr>
<tr>
<td>pSK649</td>
<td>1.6-kb EcoRI fragment carrying AVR-Pita2 from G-213 in EcoRI site of pGEM3Zf</td>
</tr>
<tr>
<td>pSK650</td>
<td>5.2-kb BanHI-EcoRI fragment carrying AVR-Pita1 from G-223 in BanHI site of pGEM3Zf</td>
</tr>
<tr>
<td>pSK923</td>
<td>2.5-kb EcoRI fragment carrying AVR-Pita3 from G-17 in EcoRI site of pGEM3Zf</td>
</tr>
<tr>
<td>pSK924</td>
<td>8.1-kb BanHI-EcoRI fragment carrying AVR-Pita3 from G-22 in BanHI-EcoRI sites of pGEM3Zf</td>
</tr>
<tr>
<td>pSK1059</td>
<td>0.9-kb Polymerase chain reaction (PCR) product from O-137 (Primers LF9 and Pita1) in pGEM-T Easy</td>
</tr>
<tr>
<td>pSK1183</td>
<td>1.0-kb PCR product from G-213 (Primers Fun1 and Fun2-2) in pGEM-T Easy</td>
</tr>
<tr>
<td>pSK1184</td>
<td>1.7-kb EcoRI-I-RHindIII fragment carrying PWR&lt;sub&gt;WR&lt;/sub&gt;:C&lt;sub&gt;WR&lt;/sub&gt;-Pua&lt;sub&gt;C&lt;/sub&gt;:T in EcoRI-I-HindIII sites of pBH2</td>
</tr>
<tr>
<td>pSK1185</td>
<td>1.7-kb EcoRI-I-RHindIII fragment carrying PWR&lt;sub&gt;WR&lt;/sub&gt;:C&lt;sub&gt;WR&lt;/sub&gt;-Pua&lt;sub&gt;C&lt;/sub&gt;:T in EcoRI-I-HindIII sites of pBH2</td>
</tr>
<tr>
<td>pSK1186</td>
<td>1.9-kb EcoRI-I-RHindIII fragment carrying PWR&lt;sub&gt;WR&lt;/sub&gt;:C&lt;sub&gt;WR&lt;/sub&gt;-Pua&lt;sub&gt;C&lt;/sub&gt;:T in EcoRI-I-HindIII sites of pBH2</td>
</tr>
<tr>
<td>pSK1187</td>
<td>1.9-kb EcoRI-I-RHindIII fragment carrying PWR&lt;sub&gt;WR&lt;/sub&gt;:C&lt;sub&gt;WR&lt;/sub&gt;-Pua&lt;sub&gt;C&lt;/sub&gt;:T in EcoRI-I-HindIII sites of pBH2</td>
</tr>
<tr>
<td>pSK1188</td>
<td>1.8-kb EcoRI-I-RHindIII fragment carrying PWR&lt;sub&gt;WR&lt;/sub&gt;:C&lt;sub&gt;WR&lt;/sub&gt;-Pua&lt;sub&gt;C&lt;/sub&gt;:T in EcoRI-I-HindIII sites of pBH2</td>
</tr>
<tr>
<td>pSK1189</td>
<td>1.8-kb EcoRI-I-RHindIII fragment carrying PWR&lt;sub&gt;WR&lt;/sub&gt;:C&lt;sub&gt;WR&lt;/sub&gt;-Pua&lt;sub&gt;C&lt;/sub&gt;:T in EcoRI-I-HindIII sites of pBH2</td>
</tr>
<tr>
<td>pSK1190</td>
<td>1.8-kb EcoRI-I-RHindIII fragment carrying PWR&lt;sub&gt;WR&lt;/sub&gt;:C&lt;sub&gt;WR&lt;/sub&gt;-Pua&lt;sub&gt;C&lt;/sub&gt;:T in EcoRI-I-HindIII sites of pBH2</td>
</tr>
<tr>
<td>pSK1191</td>
<td>1.8-kb EcoRI-I-RHindIII fragment carrying PWR&lt;sub&gt;WR&lt;/sub&gt;:C&lt;sub&gt;WR&lt;/sub&gt;-Pua&lt;sub&gt;C&lt;/sub&gt;:T in EcoRI-I-HindIII sites of pBH2</td>
</tr>
<tr>
<td>pSK2544</td>
<td>1.3-kb PCR product from 70-15 (Primers MGI0927-F and MGI0927-R) in pGEM-T Easy</td>
</tr>
</tbody>
</table>
Construction of chimeric AVR-Pita genes and functional analysis.

The primers used to amplify the promoters and coding sequences of the AVR-Pita1, AVR-Pita2, and AVR-Pita3 genes are shown in Table 3. Each primer contains a restriction enzyme site at the 5′ end to facilitate subsequent construction of chimeric genes in a modular structure: EcoRI and BamHI sites at the 5′ and 3′ ends, respectively, of the promoter, and BamHI and SplI sites at the 5′ and 3′ ends, respectively, of the coding sequences. The location of putative promoter sequences was defined relative to the translation start site (A of ATG as +1). The promoter fragments for the AVR-Pita1, AVR-Pita2, and AVR-Pita3 genes included the sequences from –9 to –489 (481 bp), –9 to –650 (642 bp), and –9 to –573 (565 bp), respectively. The last six nucleotides at the 3′ ends of the promoters were altered from 5′-TTATTT-3′ to 5′-GGATCC (BamHI site)-3′. PCR amplifications were carried out using the FailSafe PCR system. The PCR cycling program included an initial denaturation for 5 min at 95°C; two cycles of 30 s of denaturation at 95°C, 30 s of annealing at 44°C, and 1 min of extension at 72°C; and 25 cycles of 30 s of denaturation at 95°C, 30 s of annealing at 56°C, and 1 min of extension at 72°C; followed by a final extension for 10 min at 72°C. PCR products were isolated from gels using QIAquick spin columns and were cloned in pGEM-T Easy (Promega Corp.). All clones were verified by sequencing. Nine combinatorial chimeric constructs were generated by cloning a promoter (EcoRI-BamHI fragment) and a coding sequence (BamHI-SplI fragment) between the EcoRI-SplI sites of pSΔ123, which contains the N. crassa β-tubulin terminator between the SplI and HindIII sites of pGEM-3ZI (Promega Corp.). Subsequently, individual chimeric genes were isolated as EcoRI-HindIII fragments; clones containing the AVR-Pita1 or AVR-Pita2 promoters, which contain a HindIII site in the promoter, were partially digested with HindIII following a complete EcoRI digestion. The EcoRI-HindIII fragments were cloned between the EcoRI and HindIII sites of binary vector pBH2 (Mullins et al. 2001). Transformation of CP987 with these clones was conducted using Agrobacterium tumefaciens as previously described (Rho et al. 2001).

LITERATURE CITED


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


cation with subsequent expansion of pandemic clones on rice and weeds of rice. Genetics 170:613-630.


Yang, Y., Yuan, Q., and Gabriel, D. W. 1996. Water soaking function(s) of XcmH1005 are redundantly encoded by members of the *Xanthomonas avrPth* gene family. Mol. Plant-Microbe Interact. 9:105-113.


**AUTHOR-RECOMMENDED INTERNET RESOURCES**

Broad Institute: www.broad.mit.edu

Broad Institute’s Fungal Genome Initiative: www.broad.mit.edu/annotation/fungi/fgi

Broad Institute’s *Magnaporthe grisea* genetic maps website: www.broad.mit.edu/annotation/fungi/magnaporthe/maps.html

EMBL-EBI InterProScan website: www.ebi.ac.uk/InterProScan

Center for Biological Sequence Analysis SignalP 3.0 Server: www.cbs.dtu.dk/services/SignalP

National Center for Biotechnology Information’s Blast server: www.ncbi.nlm.nih.gov/BLAST