MPMI Vol. 15, No. 1, 2002, pp. 6–16. Publication no. M-2001-1116-01R.

Analysis of the Structure of the *AVR1-CO39* Avirulence Locus in Virulent Rice-Infecting Isolates of *Magnaporthe grisea*

Mark L. Farman,[1] Yukiko Eto,[3] Tomomi Nakao,[3] Yukio Tosa,[3] Hitoshi Nakayashiki,[3] Shigeyuki Mayama,[3] and Sally A. Leong[1,2]

[1]Department of Plant Pathology and [2]USDA-ARS Plant Disease Resistance Research Unit, University of Wisconsin, Madison 53706, U.S.A.; [3]Faculty of Agriculture, Kobe University, Nada, Kobe 657-8501, Japan

Submitted 5 January 2001. Accepted 8 August 2001.

The *AVR1-CO39* gene that came from a *Magnaporthe grisea* isolate from weeping lovegrass controls avirulence on the rice cultivar CO39. *AVR1-CO39* was not present in the genome of the rice-infecting *M. grisea* isolate Guy11 from French Guyana, suggesting that the gene had been deleted. Molecular analysis of the deletion breakpoints in the *AVR1-CO39* locus revealed the presence of a truncated copy of a previously unknown retrotransposon at the left-hand border. At the right-hand border was a truncated copy of another repetitive element that is present at multiple locations in the genome of Guy11. The structures of *avr1-CO39* loci were further examined in 45 rice-infecting isolates collected in Brazil, China, Japan, India, Indonesia, Mali, and the Philippines. Most isolates showed no hybridization signal with the *AVR1-CO39* probe and had the same locus structure as Guy11. Some isolates from Japan showed a signal with the *AVR1-CO39* probe, but the region specifying avirulence activity was rearranged. These findings suggest that widespread virulence to 'CO39' among rice-infecting *M. grisea* isolates is due to ancestral rearrangements at the *AVR1-CO39* locus that may have occurred early in the evolution of pathogenicity to rice.

The *AVR1-CO39* gene specifying avirulence to rice cultivar CO39 was first identified in *Magnaporthe grisea* strain 4091-5-8 (Valent et al. 1991) and subsequently in strain 2539 (Smith and Leong 1994). Both strains were developed in laboratory breeding programs. Strain 4091-5-8 inherited *AVR1–CO39* from strain K76-79, an ancestor that is pathogenic to weeping lovegrass (Valent et al. 1986), and strain 2539 inherited *AVR1-CO39* from strain 4091-5-8 (Farman and Leong 1998; Leung et al. 1988). The gene was mapped on chromosome 1 (Smith and Leong 1994) and was cloned by chromosome walking from linked restriction fragment length polymorphism markers (Farman and Leong 1998).

Corresponding author: Sally A. Leong; Telephone: +1 608-262-5309; Fax: +1 608-262-1541; e-mail: sal@plantpath.wisc.edu

Current address for Mark L. Farman: Department of Plant Pathology, University of Kentucky, Lexington, 40546, U.S.A.

M. L. Farman and Y. Eto deserve equal credit.

Subcloning and transformation experiments were used to localize the *AVR1-CO39* avirulence function to a 1.05-kb region of DNA containing a number of potential open reading frames (ORFs). However, the exact location of the *AVR1–CO39* transcriptional unit has not yet been experimentally defined. Although several lines of evidence point to one particular ORF as being the functional unit, in the current study, we define the *AVR1-CO39* avirulence gene as the minimal (1.05 kb) fragment that conferred avirulence in transformation assays.

Knowledge about molecular evolution of avirulence gene loci is of critical importance for understanding how phytopathogenic microorganisms overcome resistance in host plants. Most rice-infecting isolates of *M. grisea* are virulent on rice cultivar CO39 (Mackill and Bonman 1992), making it useful for inoculation studies as a susceptible check cultivar (Bonman et al. 1989; Roumen 1992). This observation suggests that the corresponding resistance gene recognizing *AVR1-CO39*, *Pi-CO39(t)*, is no longer effective against this host-specialized subpopulation of the fungus. To understand better why rice-infecting isolates of *M. grisea* are virulent on 'CO39', we analyzed the structure of the *avr1-CO39* locus in the virulent field isolate Guy11 (Leung and Taga 1988) as well as in other rice isolates from various countries. In all cases studied, complex genomic rearrangements were identified, each of which resulted in nonfunctional alleles.

RESULTS

Characterization of the *avr1-CO39* locus in *M. grisea* isolate Guy11.

Approximately 20 kb of the *AVR1-CO39* locus in the laboratory strain 2539 (Table 1) appears to be uncloneable in a cosmid vector, because repeated screening of a library to a depth of more than 100 genome equivalents failed to identify any clones that extended further into the locus than cosmids VIII-1 and cos18O3 (Farman and Leong 1998). However, transformation of cos18O3 into the virulent *M. grisea* isolate Guy11 resulted in a gain of avirulence to rice cultivar CO39. Subcloning of cos18O3, followed by transformation of Guy11, resulted in the localization of avirulence activity to a 1.05-kb restriction fragment that was 3 kb from the end of the insert nearest to the "uncloneable" DNA (Farman and Leong 1998). The 1.05-kb fragment containing *AVR1-CO39* did not hybridize to the genomic DNA of Guy11 (Fig. 1) and, therefore, it was necessary to use an adjacent *Bam*HI fragment (probe A) for identification of the corresponding locus in a cosmid library of Guy11 ge-

nomic DNA. This resulted in the isolation of one cosmid clone (G27-9-E) out of approximately 3,000 colonies (three genome equivalents) that were screened.

To determine if this clone spanned the deletion, Southern hybridization analysis of restricted G27-9-E DNA was performed using the 5.0-kb *Bam*HI fragment (probe A, Fig. 2A) and a 3.3-kb *Apa*I to *Not*I fragment from the *AVR1-CO39* proximal end of cosVIII-1 (probe B, Fig. 2A). As shown in Figure 2B, probe B from cosVIII-1 did hybridize to G27-9-E, confirming that its insert spanned the deletion. Furthermore, both probes hybridized to the same *Bam*HI, *Bgl*II, *Pst*I, *Sal*I, *Sma*I, and *Stu*I restriction fragments in G27-9-E (Fig. 2B), suggesting that each of these fragments encompassed the deletion. These same restriction fragments also hybridized strongly with labeled Guy11 genomic DNA (Fig. 2B), indicating the presence of repetitive DNA near the deletion breakpoints. The sizes of most of the cohybridizing fragments corresponded to those measured in Southern hybridization analyses of Guy11 genomic DNA (data not shown), indicating that the cosmid insert was truly representative of the genomic locus and had not suffered rearrangement in *Escherichia coli*.

Preliminary experiments indicated that one of the breakpoints of the deletion in the Guy11 locus was close to the priming site for primer P318 (Fig. 2A). Therefore, this primer was used to initiate the first step of a sequence "walk" using cosmid G27-9-E as a template. The sequences of the Guy11 and 2539 loci diverged 141 bp downstream from primer P318, thus marking the position of the right-hand breakpoint (Fig. 3).

DNA of cosmid G27-9-E was then used as a template to amplify the left-hand deletion breakpoint by polymerase chain reaction (PCR). The primers used were P368, which primes near the *AVR1-CO39* proximal end of cosmid VIII-1 (Fig. 2A), and P367, which primes in "novel" DNA found between the deletion borders in the Guy11 locus (shown below). PCR resulted in amplification of a 1.8-kb amplicon spanning the left-hand breakpoint. The amplicon was cloned into the *Eco*RV site of pBSKS II$^+$ for sequence analysis.

Table 1. Structures at the *AVR1-CO39* locus in *Pyricularia* isolates from rice

Isolate	Locality	Signal with CO39LR probe[a]	*AVR1-CO39* locus structure[b]	Reference or collector
2539	Lab strain	+	W	Leung et al. 1988
Guy11	Guyana	–	G1	Leung et al. 1988
ML33	Mali	–	n.d.	M. Diaby
Br10	Brazil	–	G2	S. Igarashi
Br11	Brazil	–	G2	S. Igarashi
Br12	Brazil	–	G1	S. Igarashi
Br13	Brazil	–	G2	S. Igarashi
Br15	Brazil	–	G1	S. Igarashi
Br18	Brazil	–	G2	S. Igarashi
Br21	Brazil	–	G2	S. Igarashi
Br22	Brazil	–	G2	S. Igarashi
Br26	Brazil	–	G1	S. Igarashi
Br42	Brazil	–	G1	S. Igarashi
Br127.12	Brazil	–	G1	A. S. Urashima
CH104-3	China	–	G1	Kolmer and Ellingboe 1988
CH40-1	China	–	G1	Kolmer and Ellingboe 1988
CHNOS59-6-11	China	–	G1	N. Hayashi
CHNOS60-8-1	China	–	G2	N. Hayashi
UN05-1	China	–	G1	H. Kato
UN05-2	China	–	G1	H. Kato
CHO68-138	China	–	n.d.	H. Kato
Ishi-5-1	China	–	G2	H. Kato
O-135	China	–	G1	Valent et al. 1991
0124-3	Japan	–	G2	M. Yamada
0401-4	Japan	–	G2	M. Yamada
1106-2	Japan	+	J	M. Yamada
1117-4	Japan	+	J	M. Yamada
1601-3	Japan	+	J	M. Yamada
1836-3	Japan	–	G2	M. Yamada
2012-1	Japan	+	J	M. Yamada
2303-1	Japan	+	J	M. Yamada
2403-1	Japan	–	G2	M. Yamada
4203-1	Japan	–	G2	M. Yamada
4411-2	Japan	–	G2	M. Yamada
4603-4	Japan	–	G2	M. Yamada
88A	Japan	–	G2	M. Yamada
B-157	India	–	G1	Kachroo et al. 1994
PO-02-7303	Indonesia	–	G1	A. Mukelar
PO-02-7501	Indonesia	–	G1	A. Mukelar
PO-04-7501	Indonesia	–	G1	A. Mukelar
PO-12-7402	Indonesia	–	G1	A. Mukelar
PO-12-7301-2	Indonesia	–	G1	A. Mukelar
R88130	Philippines	–	G1	Borromeo et al. 1993
R88374	Philippines	–	G1	Borromeo et al. 1993
91A58	United States	–	n.d.	A. Marchetti

[a] + = present; – = absent.

[b] Southern hybridization analysis of the *AVR1-CO39* locus in 2539 and selected *Magnaporthe grisea* isolates from rice. Genomic DNA samples were digested with *Bam*HI, electrophoresed in 0.7% agarose gel, blotted, and hybridized with the 1.05-kb fragment containing *AVR1-CO39*. nd = not determined.

Nucleotide sequence analysis of the Guy11 *avr1-CO39* locus.

The sequence of the 5′ end of the Guy11 locus was identical to the corresponding sequence in the 2539 genome contained in cosVIII-1 (Fig. 3). A 188-bp region present at the 5′ end of both the Guy11 and 2539 loci showed similarity to a repetitive element found upstream of the *M. grisea Pwl4* host specificity gene (Kang et al. 1995). Interestingly, all 36 base substitutions between the aligned sequences were either G to A or C to T transitions (data not shown). This pattern of mutation is reminiscent of the ribosome-inactivating proteins process first discovered in *Neurospora crassa* (Cambareri et al. 1989). The sequences of the Guy11 and 2539 loci diverge 210 bp downstream of the *Pwl4* element, marking the boundary of the left-hand deletion breakpoint (Fig. 3). Physical mapping experiments indicated that the sequences bordering the deleted region are approximately 20 kb apart in the 2539 genome (data not shown). However, in the Guy11 locus they are separated by just 593 bp comprising sequences not found in the *AVR1-CO39* locus of 2539. These experiments confirmed that the entire fragment containing *AVR1-CO39*, as well as approximately 19 kb of surrounding DNA, has been deleted from the corresponding *avr1-CO39* locus of Guy11.

The alternate possibility, that *AVR1-CO39* was acquired by 2539, is doubtful and will be discussed later. Consequently, we refer to the lack of the gene in Guy11 as the result of a deletion. The wild-type locus structure in 2539 is designated as W type (after its donor, a weeping lovegrass isolate) (Farman and Leong 1998). The locus in Guy11 is referred to as the G1 type (after the Guyana origin of this isolate).

Two separate repetitive elements define the left- and right-hand borders of the *AVR1-CO39* deletion.

As indicated above, the analysis of the Guy11 *avr1-CO39* locus revealed 593 bp of nucleotide sequence that was not found on endclones from either cosmid cos18O3 or cosmid cosVIII-1. Southern hybridization analysis of a PCR amplicon containing part of the 593-bp region was performed using labeled total genomic DNA from Guy11 as a probe. This revealed that at least part of the amplicon was composed of repetitive DNA (data not shown). However, the sequence of this region did not have any similarity with any of the repetitive elements already characterized in *M. grisea*.

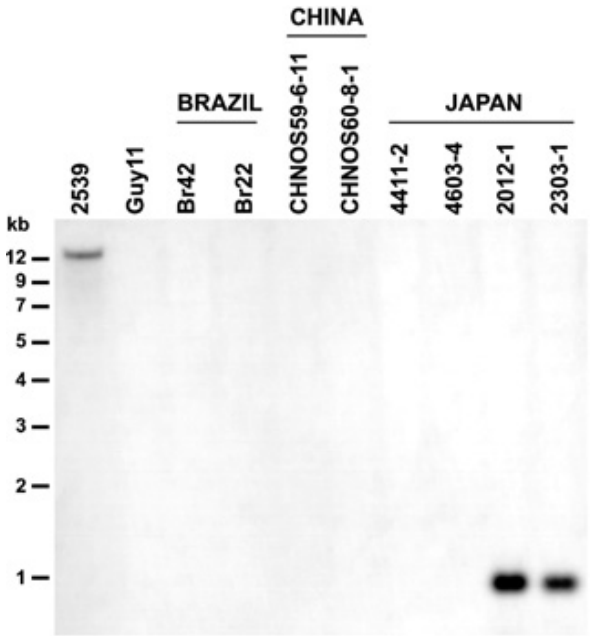

Fig. 1. Southern hybridization analysis of the *AVR1-CO39* locus in strain 2539 and selected *Magnaporthe grisea* isolates from rice**.** Genomic DNA samples were digested with *Bam*HI, electrophoresed in 0.7% agarose gel, blotted, and hybridized with the 1.05-kb fragment containing *AVR1-CO39*. Autoradiogram of the blot is shown.

To determine if the repetitive element at the G1 locus was full length, the PCR amplicon was used to isolate five additional independent copies of the element from the cosmid library of Guy11 DNA. The sequences of these elements were compared with the element identified at the Guy11 *avr* locus. Reading left from within the repetitive DNA region, all six sequences diverged at nucleotide position 471 (Fig. 3), marking one terminus of the element. This terminus coincided with the left-hand breakpoint of the deletion in the Guy11 *avr1-CO39* locus (Fig. 3). By contrast, the sequence reading to the right revealed that the various element copies were rather heterogeneous in their structures. The element at the Guy11 *avr1-CO39* locus possesses 2.5 tandem repeats of a 45-bp sequence (Fig. 3), but in the other copies of the element that were examined, this sequence was repeated only 1.5 times. In three of the six copies, this motif defined the 5′ terminus of the insertion, as their sequences diverged beyond this point. However, in the three remaining copies, the sequence identity continued beyond this region, indicating that elements terminating at the direct repeats were truncated (data not shown). The regions sequenced were not similar to any entries in either the GenBank database or the North Carolina State University (NCSU) *M. grisea* BAC-end database. However, in an unrelated study of isolates that attack *Lolium perenne* (M. L. Farman, *unpublished data*), a longer representative of the same element was identified. Part of it showed a highly significant (*P* values between 1×10^{-11} and 1×10^{-17}) similarity to reverse transcriptase domains of copialike retrotransposons from *Arabidopsis thaliana* (Henikoff and Comai 1998), *Nicotiana tabacum* (Grandbastien et al. 1989), and wild rice *Oryza australiensis* (Noma et al. 1997). We conclude that the element residing at the deletion breakpoint of the Guy11 *avr1-CO39* locus is part of a new *M. grisea* retrotransposon, which we have called RETRO5.

The distribution of RETRO5 in various host-specialized populations of *M. grisea* was determined by using a 3.7-kb *Pst*I fragment from within the element to probe genomic DNA of *M. grisea* isolates from various hosts (Table 2). This analysis revealed that RETRO5 was widespread in *M. grisea*, with the highest copy numbers occurring in isolates from *Setaria* spp. (Fig. 4). One isolate from crabgrass lacked the element (lane 2) and isolates from St. Augustine grass (lanes 7 and 8) had very few copies.

There were 188 bp of non-RETRO5 sequence between the truncated RETRO5 element and the right-hand border of the deletion. The NCSU *M. grisea* BAC-end database was searched with this sequence and numerous BAC-end sequences were returned, some of which were identical to the entire 188-bp sequence. The presence of many similar sequences in the database indicates that this sequence is part of another repetitive element we refer to as REP1. Alignment of REP1 with the sequences returned by the database search revealed that one terminus of REP1 coincided with the right-hand deletion breakpoint (Fig. 3). However, the REP1 copy present in the *avr1-CO39* locus is also truncated, having been interrupted by the RETRO5 sequence (Fig. 3).

Structure of the *avr1-CO39* loci in other *M. grisea* isolates from rice.

The rice cultivar CO39 is almost universally susceptible to *M. grisea* isolates from rice, leading us to hypothesize that these isolates constitute a population lacking the intact *AVR1-*

CO39 gene. To test this hypothesis, genomic DNA of 45 rice-infecting isolates from Brazil, China, Japan, India, Indonesia, Mali, the Philippines, and the United States were probed with the 1.05-kb fragment containing *AVR1-CO39* (CO39LR probe). Only five isolates, all from Japan, yielded a strong signal with CO39LR. The remainder showed no signals, indicating they lack the gene entirely (Fig. 1, Table 1). Inoculation studies showed that the Japanese isolates were highly virulent

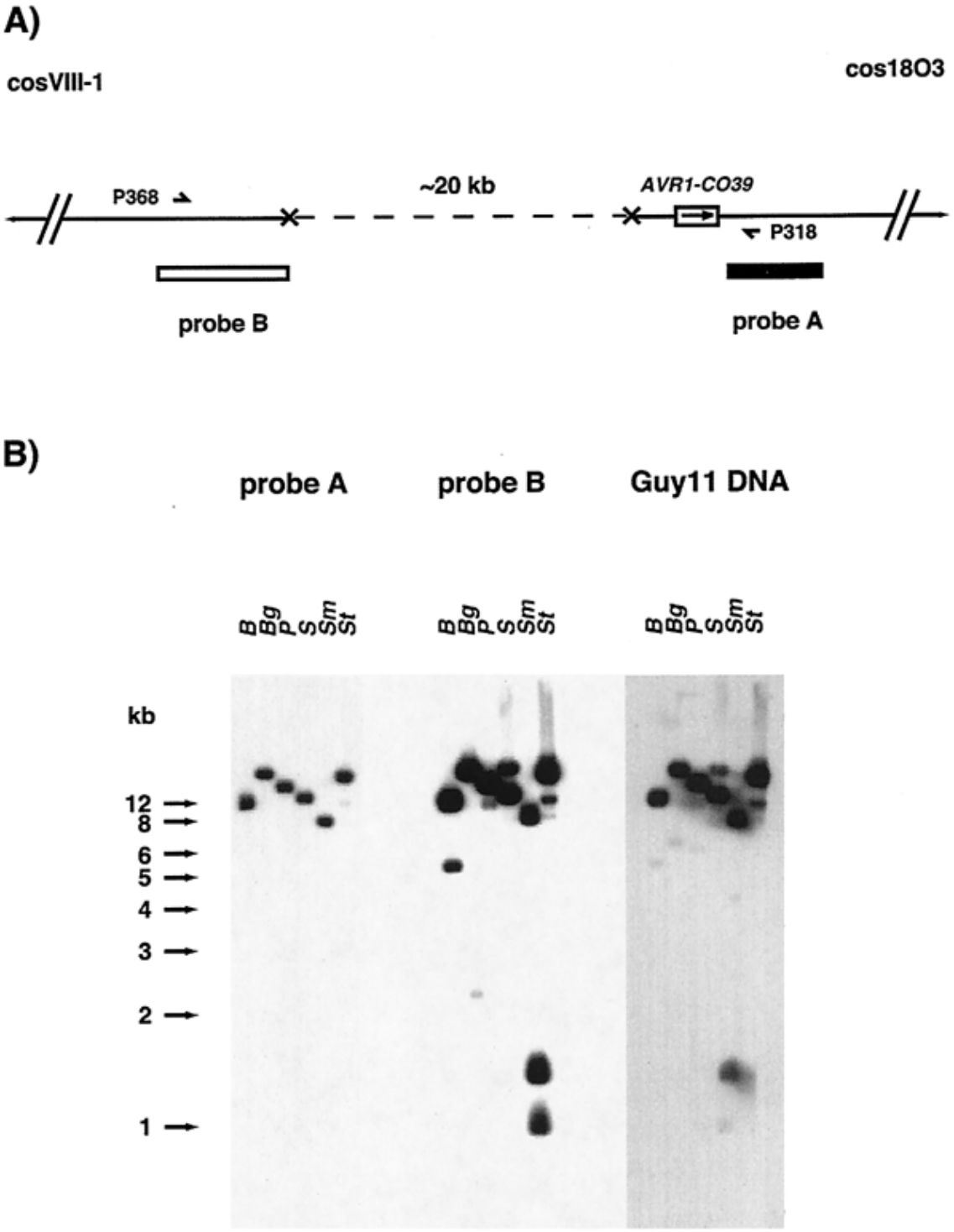

Fig. 2. Isolation and characterization of a cosmid carrying the Guy11 *avr1-CO39* locus. **A,** Organization of the *AVR1-CO39* locus in strain 2539. Approximately 20 kb of DNA linked to the avirulence gene was not cloneable in a cosmid vector. This is represented with a dashed line and is delimited by X's. DNA flanking the missing region was present in cosmids cosVIII-1 and cos18O3 containing inserts derived from genomic DNA of strain 2539 (Farman and Leong 1998). Probe A identified cosmid G27-9-E in a cosmid library of Guy11 genomic DNA (M. Farman and S. A. Leong, *unpublished data*). Also shown are locations of primers that were used to amplify DNA from the Guy11 locus. **B,** Cosmid G27-9-E was restricted with various enzymes and electrophoresed in an agarose gel and then blotted. The blot was hybridized sequentially with a 5.0-kb *Bam*HI fragment from cos18O3 (probe A); a 3.3-kb *Apa*I-to-*Not*I fragment from cosVIII-1 (probe B). Finally, genomic DNA of Guy11 was labeled and used to probe G27-9-E for the purpose of identifying repetitive DNA fragments. Repetitive DNA fragments hybridized strongly to the probe. B = *Bam*HI, Bg = *Bgl*II, P = *Pst*I, S = *Sac*I, Sm = *Sma*I, and St = *Stu*I.

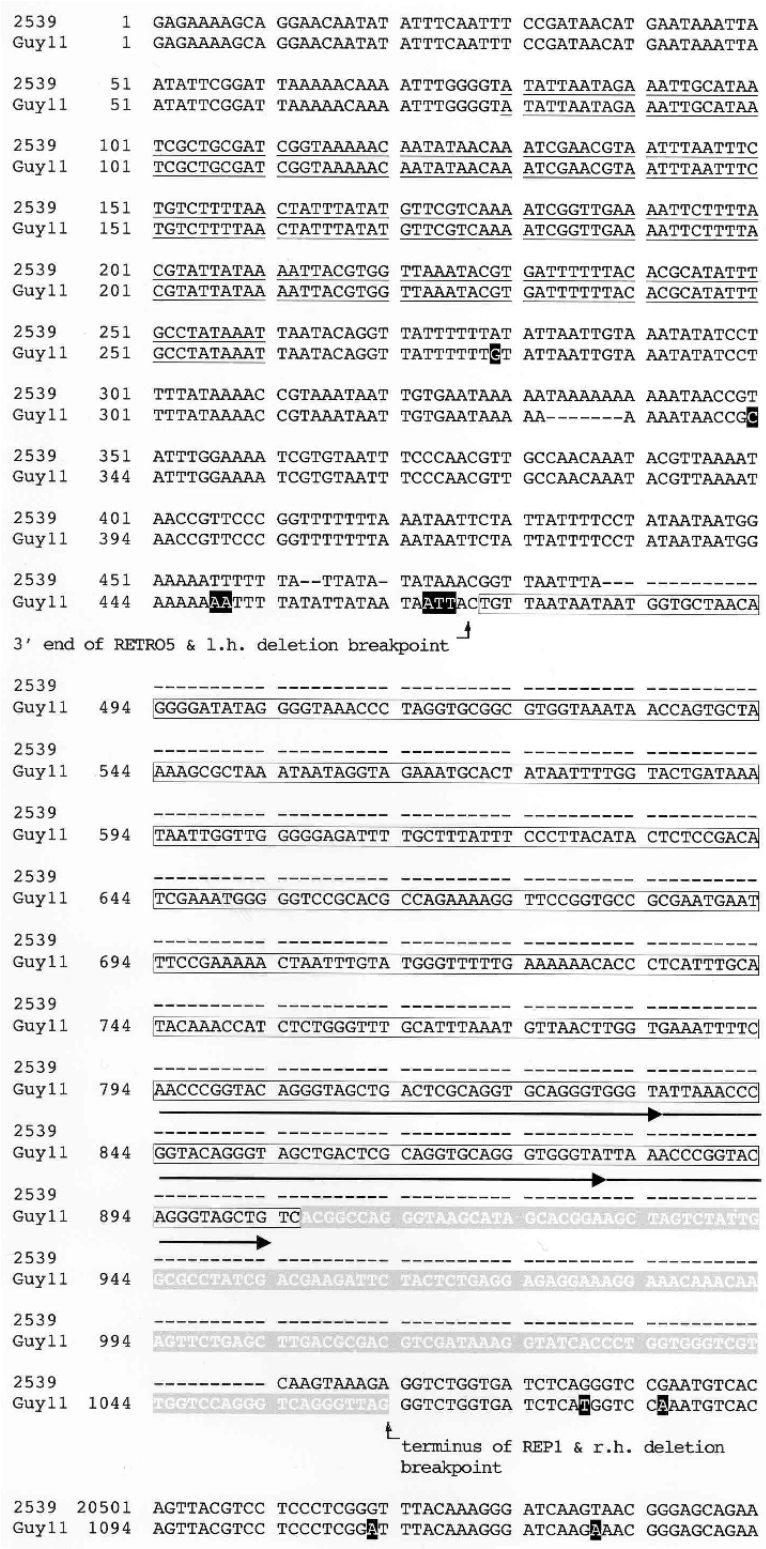

Fig. 3. Nucleotide sequence alignment between the *avr1-CO39* locus of Guy11 and sequences on either side of the *AVR1-CO39* gene of strain 2539. The dotted line in the middle of the 2539 sequence represents approximately 20 kb of DNA that is not present in the *avr1-CO39* locus of Guy11. The nucleotide positions at the 3′ end of the 2539 sequence are approximate. Sequence shown at the 5′ end of the dotted line was present in cosmid cosVIII-1, and the sequence that is 3′ was in cos18O3 (Fig. 2). The underlined sequence is a truncated copy of the *Pwl4*-associated repetitive element that exhibited ribosome-inactivating proteinlike mutations. 'Novel' sequences not found in the *AVR1-CO39* locus of 2539 are shown as boxed text. The region that is part of the RETRO5 element is shown as plain text in an unshaded box. Tandem repeats of 45 bp that mark the right-hand border of RETRO5 are underlined with arrows. The sequence from the truncated REP1 element is shown as white text in a shaded box. Vertical arrows indicate deletion breakpoints and the termini of RETRO5 and REP1. r.h. = right hand, l.h. = left hand.

on rice cultivar CO39 (Table 1). Therefore, the *AVR1-CO39* homologs present in these isolates do not function as avirulence genes, at least on this cultivar.

Inverse PCR was used to isolate the *AVR1-CO39* homolog and flanking DNA from genomic DNA of one of the Japanese isolates (2012-1), and the amplicon was cloned in the pBluescript vector. Nucleotide sequence analysis revealed that the sequences corresponding to the left half of the 1.05-kb fragment that contains *AVR1-CO39* were missing. To confirm this result, the membrane shown in Figure 1 was reprobed with the right and left halves of the CO39LR probe. As expected, the right-half probe (CO39R) produced the same signal as the full-length probe (CO39LR), whereas the probe for the left half (CO39L) produced no signals (data not shown). This demonstrated that the sequences corresponding to the left half of the probe are not present in the genome of isolate 2012-1.

The sequence of the amplicon obtained from 2012-1 revealed that, as in Guy11, the right-hand border of the deletion in this strain is also defined by the terminus of a REP1 element (Fig. 5). However, in the G1 locus, the REP1 element present at the deletion border was interrupted by a portion of the RETRO5 element, whereas the copy of REP1 at the *avr1–CO39* locus of 2012-1 was truncated by the repetitive elements MGR608 and MGR619 (Kang et al. 1995) (Fig. 5). To the left of MGR608 and MGR619 were sequences from another RETRO5 element. These sequences were derived from a region of RETRO5 that was different from the one present in the G1 locus (data not shown). In addition, the copy of RETRO5 in the *avr1-CO39* locus of 2012-1 was interrupted by insertion of a MAGGY retrotransposon (Farman et al. 1996) (Fig. 5). PCR with primers for the region to the left of MGR608 and MGR619 and to the right of *AVR1-CO39* showed the four other Japanese isolates that gave a hybridization signal with CO39LR all have the same locus structure as 2012-1. Additional confirmation was obtained by Southern hybridization analyses using probes from the left-hand side of the locus and the CO39R probe (data not shown). Hereafter, we refer to the locus structure in the Japanese isolates as J type.

Interestingly, the deletion in the 1.05-kb fragment failed to eliminate the ORF that is most likely to encode the *AVR1–CO39* gene product (M. L. Farman and S. A. Leong, *unpublished data*). However, its translational stop codon was mutated, leading to an amino acid extension in the predicted translation product. No alternative stop codon was found in the sequenced portion of the J-type locus.

Once the structures of the G1- and J-type deletions had been established, a PCR-based strategy was used to survey the structure of the *AVR1-CO39* locus in the rice isolate population. A total of four forward primers and two reverse primers was used, and six primer combinations were tested for each isolate (Figs. 5 and 6). This survey resulted in the discovery of another structure at the *AVR1-CO39* locus, designated G2 (Fig. 5). The PCR profile of the G2 type was the same as that of the G1 type, with five of the six primer combinations, but differed from the P368-P318 combination (Figs. 5 and 6). It was assumed that the G2 type has some rearrangement(s) in the region upstream of primer GF1 (Fig. 5). This was confirmed by PCR analyses, with additional primers designed between P368 and GF1 (data not shown).

The P368-P318 primer combination also produced unexpected weak bands in the W and J types (Fig. 6B). The origin of these bands is unknown but may be nonspecific amplicons initiated by priming in the right-hand region of the locus. The results of this survey are summarized in Table 1. We conclude that the five Japanese isolates that yielded the hybridization signal with the CO39RL probe all possess the J-type structure. Isolates lacking signals were either the G1 and G2 types, with the exception of isolates CHO68-138 (China), ML33 (Mali), and 91A58 (United States), which could not be classified into any of the four types and, presumably, had a novel locus structure (or structures).

DISCUSSION

A significant finding from the survey of *M. grisea* isolates conducted in this study was that an intact *AVR1-CO39* locus was not present in any of the isolates from rice. This provided a rational explanation why cultivar CO39 is almost universally susceptible to *M. grisea* isolates from rice. It appears that virulence of *M. grisea* to a particular cultivar or host species is

Table 2. *Magnaporthe grisea* isolates used to survey the distribution of RETRO5 by Southern hybridization analysis in Figure 4

Isolate	Host	Locality	Reference or collector
DsDc1	Crabgrass	U.S.A.	M. Farman
Ds88428	Crabgrass	Phillipines	Borromeo et al. 1993
G22	Weeping lovegrass	Japan	Kang et al. 1995
G17	Goosegrass	Japan	Kang et al. 1995
PH42	Finger millet	Philippines	D. Tharreau
RW12	Finger millet	Rwanda	D. Tharreau
St1108	St. Augustine grass	United States	L. Tredway, L. Burpee
St1204	St. Augustine grass	United States	L. Tredway, L. Burpee
Br32	Wheat	Brazil	S. Igarashi
Br80	Wheat	Brazil	S. Igarashi
LpKY96-1	Perennial ryegrass	United States	P. Vincelli
LpkY97-1	Perennial ryegrass	United States	M. Farman
LpFa97-1	Tall fescue	United States	M. Farman
Fa1207	Tall fescue	United States	L. Tredway, L. Burpee
91T32	Rice	United States	Levy et al. 1991
89L36	Rice	United States	Levy et al. 1991
Fr9	Rice	France	D. Tharreau
926-20	Rice	Philippines	M. Bernado, M. Bronson
YF1	Yellow foxtail	United States	M. Farman
YF2	Yellow foxtail	United States	M. Farman
Arc2	Green foxtail	United States	M. Farman
McC1	Green foxtail	United States	M. Farman
GF1	Giant foxtail	United States	M. Farman
GF2	Giant foxtail	United States	M. Farman
SiKY97-1	German foxtail millet	United States	P. Bachi
SiKY97-3	German foxtail millet	United States	P. Bachi

commonly associated with deletions of, and insertions within, the appropriate avirulence gene loci. Deletion of AVR and host-specificity genes has been observed at the *AVR2-YAMO* (renamed *AVR-Pita*; Jia et al. 2000) locus (Orbach et al. 2000), at *Pwl2* loci (Kang et al. 1995), and at the *AVR1-MARA* locus (Mandel et al. 1997). However, in the case of *AVR2-Pita* and *Pwl2*, nonfunctional, i.e., virulence, alleles are also commonly found (Kang et al. 1995; Orbach et al. 2000). Some virulence alleles of *AVR-Pita* have been found to contain point mutations, and one had a copy of the MGR586 (Pot3) element inserted into it (Kang et al. 2001; Orbach et al. 2000). Similarly, analysis of a virulence allele of the recently cloned *AVR-IRA7* avirulence gene revealed that a SINE element had disrupted the locus (Bohnert et al. 2000).

In *Cladosporium fulvum*, the molecular mechanisms underlying loss of avirulence function appear to depend on the avirulence gene. All of the *C. fulvum* races that are virulent to tomato cultivars carrying the *Cf9* resistance gene have deleted the corresponding avirulence gene, *Avr9* (Van den Ackerveken et al. 1992). In contrast, virulence to cultivars with the resis-

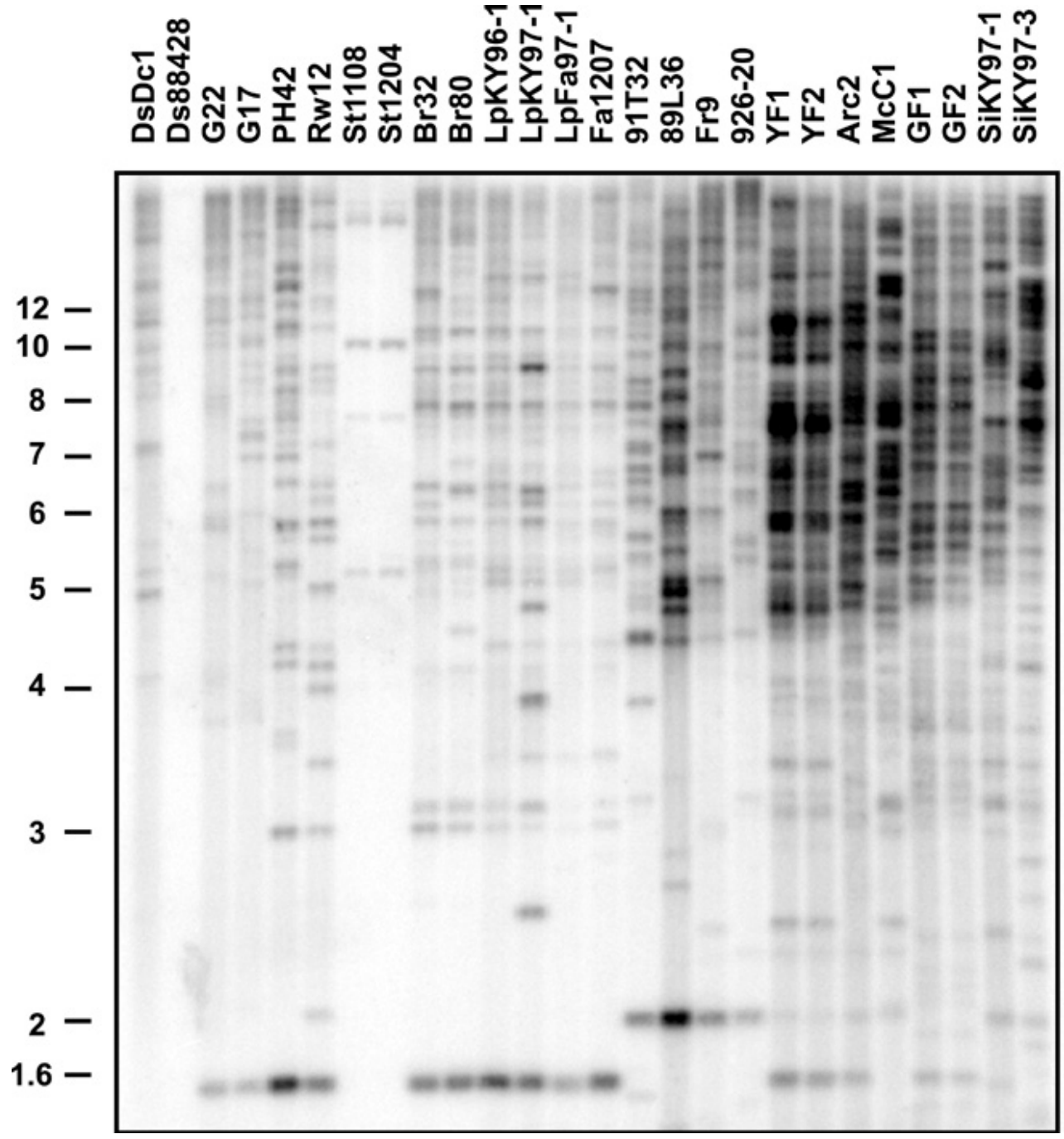

Fig. 4. Distribution of RETRO5 among various host-specific forms of *Magnaporthe grisea*. Aliquots of *Bam*HI-digested genomic DNA from each isolate were electrophoresed in a 0.7% agarose gel (200 ng per lane) that was then washed and Southern-blotted using previously described protocols (Thornbury and Farman 2000). The blot was probed with a 3.7-kb *Pst*I fragment from within the RETRO5 transposon. The phosphorimage is shown. The isolates' hosts of origin were DcDs1 and Dc88428 (crabgrass); G22 (weeping lovegrass); G17 (goosegrass); PH42 and RW12 (finger millet); St1108 and St1204 (St. Augustine grass); Br32and Br80 (wheat); LpKY96-1 and LpKY97-1 (perennial ryegrass); FaKY97-1 and Fa1207 (tall fescue); 91T32, 89L36, Fr9, and 926-20 (rice); YF1 and YF2 (yellow foxtail); Arc2 and McC1 (green foxtail); GF1 and GF2 (giant foxtail); and SiKY97-1 and SiKY97-3 (German foxtail millet). Molecular sizes in kilobases are shown on the left.

tance gene *Cf4* is due exclusively to point mutations in the *Avr4* gene (Joostens et al. 1994). In the case of the barley scald pathogen, *Rhynchosporium secalis*, virulence is associated with either a complete lack of the *NIP1* avirulence gene or possession of mutated alleles (Rohe et al. 1995).

The patterns of evolution at the *AVR1-CO39* locus are complex. The 3′ portion of the RETRO5 element present at the G1 and G2 loci is truncated at the point where REP1 inserted (Figs. 3 and 5). It is likely that the 5′ end of RETRO5 was removed when the resident REP1 element recombined with another copy downstream of *AVR1-CO39*. This event would also have caused the deletion of *AVR1-CO39*. The finding that the immediate borders of the *AVR1-CO39* deletions are identical in the G1- and G2-type loci clearly indicate that these loci were derived from the same ancestral deletion event. The G1 locus most likely represents the original deletion, with the G2 structure being the result of a subsequent rearrangement that separated the Pwl4 element from the RETRO5 and REP1 elements that reside in the G1 locus (Y. Tosa et al., *unpublished data*). With both the left- and right-hand breakpoints differing from the former loci, the deletion in the J-type locus appears to be quite different from the G types. Specifically, only REP1 is found immediately adjacent to the deletion boundary with the MGR619/608 elements. Although a portion of RETRO5 is present in the J-type locus, it is different from the region in the G1- and G2-type loci. Currently, we suspect that the involvement of REP1 in both J- and G-type deletions may be coincidental.

As mentioned above, an alternative interpretation to deletion of *AVR1-CO39* from the Guy11 genome is that the Guy11 locus actually represents the ancestral form and K76-79 (the original donor of *AVR1-CO39*), or a progenitor, acquired the gene by horizontal transfer. However, the presence of RETRO5 in the *avr1-CO39* (G type) loci strongly disfavors this alternative hypothesis. If the Guy11 locus represents the ancestral form or a derivative thereof, then acquisition of *AVR1-CO39* would have occurred through a very precise substitution of RETRO5/REP1 with a completely nonhomologous segment of DNA containing *AVR1-CO39*. This scenario is unlikely and until we have evidence to suggest otherwise, we will assume that *AVR1-CO39* was deleted.

A small number of rice-infecting *M. grisea* isolates from the Philippines are avirulent on 'CO39' (Zeigler et al. 1995). However, hybridization analysis of genomic DNA from some of these have shown they lack *AVR1-CO39*, and PCR analysis of a representative isolate revealed the G1 locus structure (M. L. Farman and S. A. Leong, *unpublished data*). It is interesting to note that a quantitative trait locus mapping study performed by Wang and associates (1994) resulted in the identification of a resistance gene in 'CO39' that is effective against certain Philippine rice-infecting isolates of *M. grisea*. However, this gene is distinct from the gene recognizing *AVR1–CO39*, *Pi-CO39(t)*, and maps at another chromosomal location (Leong et al. 2000). Based on these two observations, it appears that the Philippine isolates have another gene (*AVR2–CO39*?) that specifies avirulence to 'CO39'.

None of the rice-infecting *M. grisea* isolates we have tested to date have an intact (W type) *AVR1-CO39* locus (this study; and authors' unpublished data). Therefore, the deletion events characterized in this study may have occurred prior to the evolution of the rice-infecting form of *M. grisea* or before the divergence of the currently known rice-infecting lineages. Currently, it is unclear exactly when the deletion of *AVR1-CO39* occurred during the evolution of *M. grisea*. A PCR-based analysis of DNA from host-specific forms that are phylogenetically related to rice pathogens may provide an answer to this question.

The finding that *AVR1-CO39* is entirely absent from the genomes of most of the rice pathogens studied indicates the gene probably does not function as a virulence factor on compatible hosts. Therefore, *AVR1-CO39* appears to differ in this regard from the *C. fulvum* avirulence determinants *ECP1* and *ECP2*

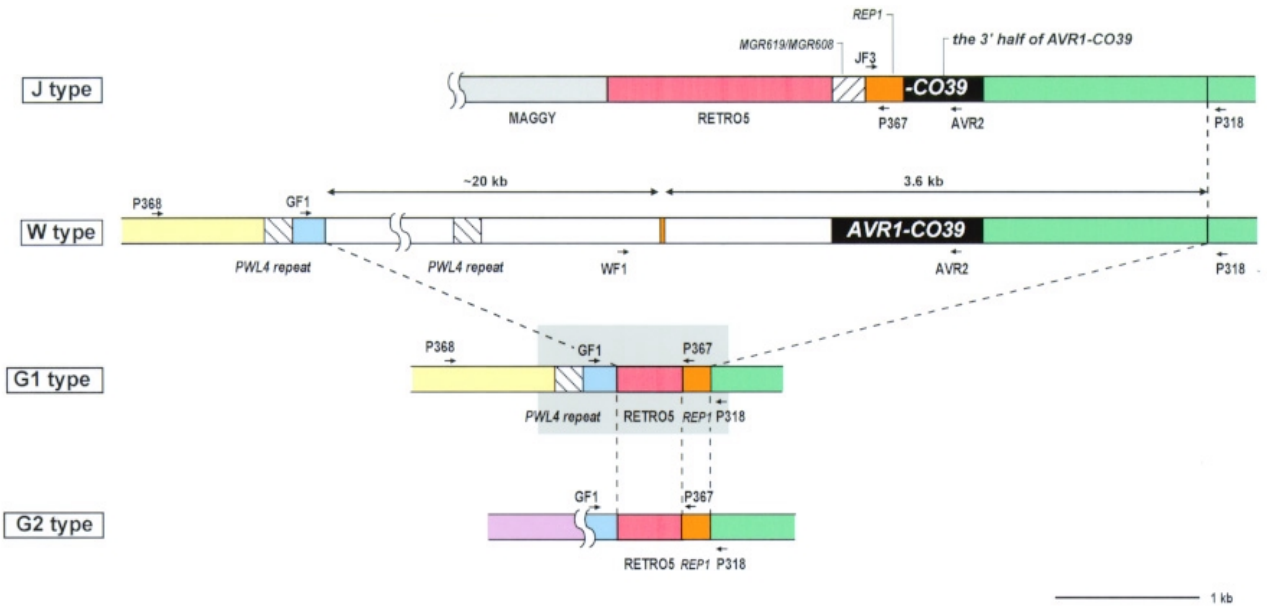

Fig. 5. Comparative maps showing the physical organization of the locus surrounding three *avr1-CO39* null alleles. Locus type is indicated on the left. The smallest restriction fragment in the W-type locus of strain 2539 that contained full avirulence activity is shown as a black box labeled with white lettering. A truncated version of this fragment, found in the J-type locus, is shown as a partially labeled black box. Chromosomal regions that have identical or nearly identical sequences are shown in the same color. Repetitive elements are shown in the following colors or shading patterns: *Pwl4*, diagonal bars; RETRO5, red; REP1, orange; MGR608/MGR819, reversed diagonal bars; and MAGGY, dark blue. Transposon names are also noted beneath or above the respective boxes. Locations of primers used in this study are indicated by small arrows showing the direction of strand synthesis. Dashed lines are included to provide reference points for places of alignments between the loci. The region of the G1-type locus, whose sequence is shown in Figure 3, is indicated with a shaded background. Scale is indicated by the bar in the bottom right-hand corner and in the body of the figure.

(Lauge et al. 1997) and the *NIP1* gene of *Rhynchosporium secalis* (Rohe et al. 1995). However, specific *AVR1-CO39* gene knockout experiments are needed before a possible role in virulence can be completely discounted.

Our results provide a clear explanation of why 'CO39' is almost universally susceptible to *M. grisea* isolates that are pathogenic to rice. The breakpoints and overall organizations of the J- and G-type loci are different, suggesting they may

A

Primers		Expected size(kb) of PCR products			
Forward	Reverse	W	G1	G2	J
① P368	P318	-	1.9	-	-
② GF1	P318	-	0.9	0.9	-
③ JF3	P318	-	-	-	2.3
④ JF3	AVR2	-	-	-	0.49
⑤ WF1	P318	4.0	-	-	-
⑥ WF1	AVR2	2.1	-	-	-

B

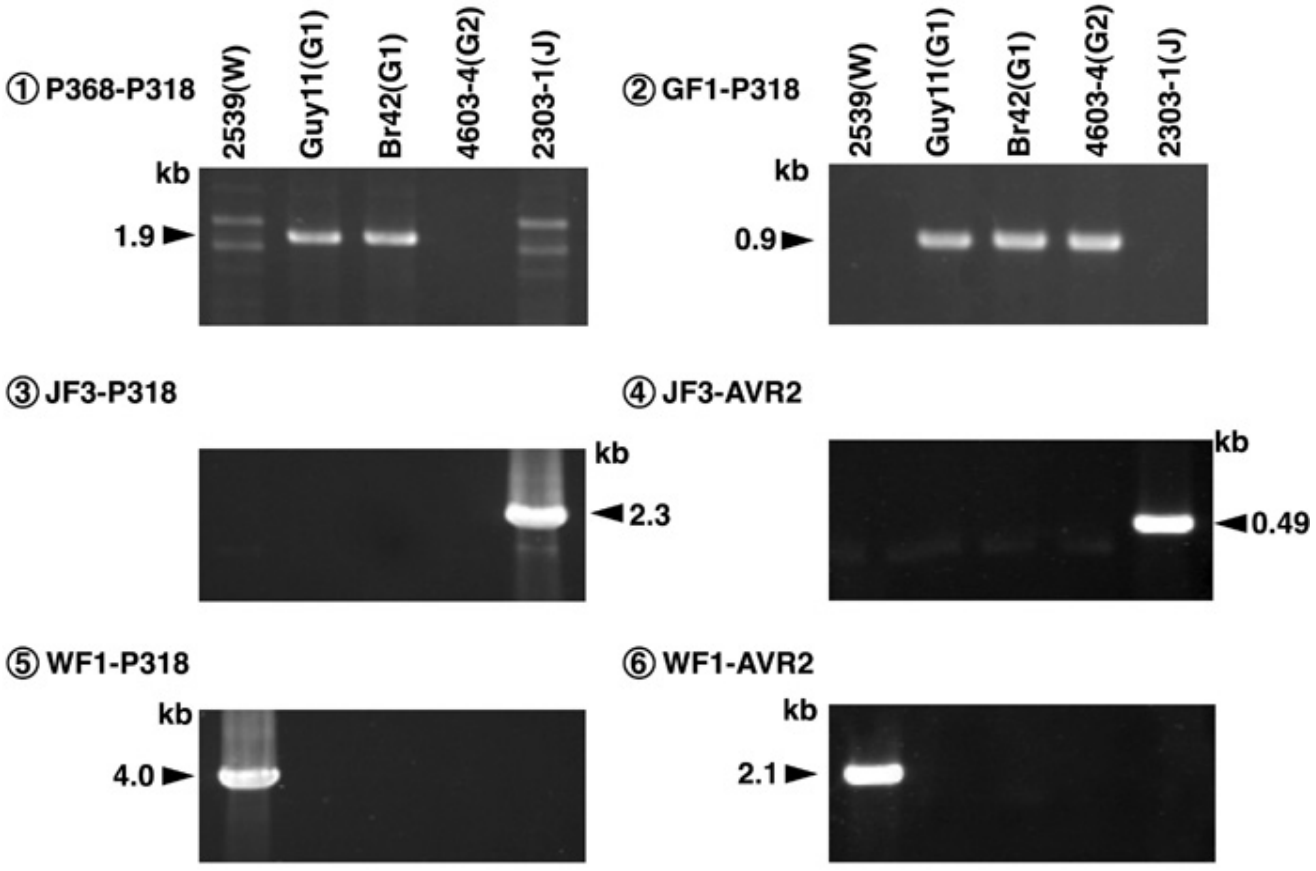

Fig. 6. Allele-specific polymerase chain reactions (PCR) of *avr1-CO39* loci from *Magnaporthe grisea* strains pathogenic to rice. Based on the physical organizations of the *avr1-CO39* loci shown in Figure 5, primer pairs were designed to enable diagnostic amplification of the different loci from various *M. grisea* isolates. These primers were then used to perform PCR on DNA of other *M. grisea* strains listed in Table 1. **A,** Primer pairs used for amplification and the expected sizes of the products in kilobase pairs. **B,** The ethidium bromide-stained agarose gels containing electrophoresed amplification products. Primers pairs used are noted next to each gel photo.

have arisen by independent deletion events. The implication from these results is that, historically, there may have been strong selection for loss of avirulence conferred by *AVR1–CO39.*

MATERIALS AND METHODS

Fungal isolates and pathogenicity tests.

All *M. grisea* isolates used in this study and their origins are listed in Tables 1 and 2. Selected isolates that possess the various *avr1-CO39* locus structures were subjected to pathogenicity tests. Inoculations were performed as described by Smith and Leong (1994).

DNA isolation and hybridization.

Fungal isolates were grown with shaking in 50 ml of complete medium broth (3 g of yeast extract, 3 g of casamino acids, and 5 g of sucrose per liter) in Erlenmeyer flasks at 23°C for 4 days. Total DNA was extracted from the resulting mycelia as described by Nakayashiki and associates (1999) or Farman and Leong (1995). Total genomic DNA was digested with a restriction enzyme and fractionated through 0.7% agarose gels in 0.5× Tris-borate-EDTA buffer. The fractionated DNA was transferred to Magnagraph membranes (MagnaGraph; Osmonics, Westborough, MA, U.S.A.) and fixed by UV irradiation following the manufacturer's instructions. The membranes were hybridized overnight using one of two methods: with biotin-labeled probes in 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) containing 5× Denhardt's solution, 0.5% sodium dodecyl sulfate (SDS) and 100 μg of sonicated salmon sperm DNA per ml at 68°C, or with ^{32}P-labeled probes, according to the protocol described by Thornbury and Farman (2000). After hybridization, the membrane was washed twice in 2× SSC, 0.1% SDS for 5 min at room temperature, and then washed twice in 0.1× SSC containing 0.1% SDS for 15 min at 68°C. Detection of target DNA sequences was performed using the Phototope Star Detection Kit (New England Biolabs, Beverly, MA, U.S.A.) or by phosphorimaging.

Cloning of the *avr1-CO39* locus of Guy11.

A 5.0-kb *Bam*HI fragment from cos18O3 spanned the right-hand deletion breakpoint and was, therefore, used to probe a cosmid library of Guy11 genomic DNA to isolate clones containing sequences flanking the deleted gene. Sequencing with primer P318 (5′ AGTTAGCACTCACTTTA 3′) was used to localize one deletion breakpoint and then sequences flanking the entire deleted *AVR1-CO39* locus were amplified using the primer pairs P367 (5′ AATAGACTAGCTTCCGT 3') and P368 (5′ AACTGCGCTGAGCTACC 3') and P318 and P368. PCR was performed using 100 pg of cosmid G27-9-E as template with the following reaction conditions: 1× Pfu buffer (Stratagene, La Jolla, CA, U.S.A.); 20 pmol of each primer, 50 μM nucleotides, and 2.5 U *Pfu* polymerase (Stratagene). Cycle parameters were 96°C for 1 min; followed by 35 cycles of 95°C for 30 s, 44°C for 30 s, 72°C for 1 min 15 s; followed by a single cycle of 5 min at 72°C.

Characterization of a repetitive DNA element at the Guy11 *avr1-CO39* locus.

The consensus structure of the repetitive element at the Guy11 *avr1-CO39* locus was determined by comparing its sequence with five homologous elements from other locations in the Guy11 genome. These additional elements were identified by screening the Guy11 cosmid library (Farman and Leong 1998) with a PCR amplicon containing repetitive DNA from within the 600-bp DNA at the Guy11 *avr1-CO39* locus. Restriction fragments containing new elements were subcloned into pBSKS II$^+$ and sequenced. The terminus of the element was defined as the point where the sequences of the different elements diverged.

Cloning of the *avr1-CO39* locus of a Japanese rice isolate carrying a homolog.

Southern hybridization analysis revealed that some Japanese isolates carry the right half of *AVR1-CO39* but lack the left half. The flanks of the right half were amplified by using inverse PCR with a set of primers, AVR1 (5′-GCGAATCCATAGACAAGGAC-3′) and WF3 (5′-TTCCTACTTTCGTCCCATTC-3′). Genomic DNA (1 μg) was digested completely with *Eco*RI, extracted with phenol/chloroform, precipitated with ethanol, and self-ligated at 16°C for 16 h in a 200-μl reaction containing 1 mM of ATP, 700 U of T4 DNA ligase (TaKaRa, Tokyo), and 1× ligation buffer (TaKaRa). The ligation product was precipitated with ethanol and resuspended in 20 μl of Tris-EDTA. Two microliters corresponding to approximately 100 ng of DNA of the resuspended DNA was added to a 50-μl reaction containing 0.4 μM of each primer, 0.4 mM of each dNTP, 2.5 U LA Taq polymerase, and LA PCR buffer with 2.5 mM $MgCl_2$ (TaKaRa) and amplified using the following conditions: 94°C for 1 min, followed by 30 cycles of 98°C for 20 s and 67°C for 10 min, with a final extension period at 72°C for 10 min. The amplicon was cloned at the *Eco*RV site of pBluescript SKII$^+$ and sequenced.

Allele-specific PCR.

The locus structure in the other rice-pathogenic isolates was characterized by allele-specific PCR. Primers used were P368, GF1 (5′-TTGGAAAATCGTGTAAT-3′), JF3 (5′-CACGGCCAGGGTAAGCATAG-3′), WF1 (5′-AAGGGAAGAGTAAATGAAATGTAGA-3′), P318, and AVR2 (5′-AAAATCGCCAGAAAGGAC-3′). The amplification was performed in a 50-μl reaction containing 1.25 U of HotStarTaq DNA polymerase (Qiagen, Hilden, Germany), 1× PCR buffer provided by the manufacturer, 200 μM of each dNTP, 0.2 μM of each primer, and 100 ng of template DNA. Reactions were heated to 95°C for 15 min and then amplified for 30 cycles (1 min at 94°C, 1 min at 44 to 55°C, depending on the primer's melting (*or* midpoint) temperature value, and 1 min at 72°C), with a final extension period at 72°C for 10 min.

ACKNOWLEDGMENTS

This work was supported by grants to S. A. Leong from the Rockefeller Foundation and the USDA-ARS and to Y. Tosa from the Japanese Ministry of Science.

LITERATURE CITED

Bohnert, H. U., Fudal, I., Dioh, W., Tharreau, D., Notteghem, J.-L., and Le Brun, M.-H. 2000. Identification of the avirulence gene *AVR-IRA7* of the rice blast fungus *Magnaporthe grisea*. Abstr. 4th Int. Rice Gen. Symp. 2000. IRRI, Los Banos, The Philippines.

Bonman, J. M., Bandong, J. M., Lee, Y. H., Lee, E. J., and Valent, B. 1989. Race-specific partial resistance to blast in temperate japonica rice cultivars. Plant Dis. 73:496-499.

Borromeo, E. S., Nelson, R. J., Bonman, J. M., and Leung, H. 1993. Genetic differentiation among isolates of *Pyricularia* infecting rice and weed hosts. Phytopathology 83:393-399.

Cambareri, E. B., Jensen, B. C., Schabtach, E., and Selker, E. U. 1989. Repeat-induced G-C to A-T mutations in *Neurospora.* Science 244:1571-1575.

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

Farman, M. L., and Leong, S. A. 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., Tosa, Y., Nitta, N., and Leong, S. A. 1996. MAGGY: A retrotransposon found in the genome of the rice blast fungus, *Mag-*

naporthe grisea. Mol. Gen. Genet. 251:665-674.
Grandbastien, M. A., Spielmann, A., and Caboche, M. 1989. Tnt1: A mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature 337:376-380.
Henikoff, S. and Comai, L. 1998. A DNA methyltransferase homolog with a chromodomain exists in multiple polymorphic forms in *Arabidopsis*. Genetics 149:307-318.
Jia, Y., McAdams, S. A., Bryan, G. T., Hershey, H. P., and Valent, B. 2000. Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. EMBO (Eur. Mol. Biol. Organ.) J. 19:4004-4014.
Joostens, M. H. A., Cozinnsen, T. J., and De Wit, P. J. G. M. 1994. Host resistance to a fungal tomato pathogen lost by a single base-pair change in an avirulence gene. Nature 367:483-386.
Kachroo, P., Leong, S. A., and Chattoo, B. 1994. Pot2, an inverted repeat transposon from the rice blast fungus *Magnaporthe grisea*. Mol. Gen. Genet. 245:339-348.
Kang, S., Sweigard, J. A., and Valent, B. 1995. The *PWL* host specificity gene family in the rice blast fungus *Magnaporthe grisea*. Mol. Plant-Microbe Interact. 8:939-948.
Kang, S., Lebrun, M. H., Farrall, L., and Valent, B. 2001. Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. Mol. Plant-Microbe Interact. 14:671-674.
Kolmer, J. A., and Ellingboe, A. H. 1988. Genetic relationships between fertility and pathogenicity and virulence to rice in *Magnaporthe grisea*. Can. J. Bot. 66:891-897.
Lauge, R., Joosten, M. H. A. J., van den Ackerveken, G. F. J. M., van den Broek, H. W. J., and de Wit, P. G. J. M. 1997. The in planta-produced extracellular proteins ECP1 and ECP2 of *Cladosporium fulvum* are virulence factors. Mol. Plant Microbe Interact. 10:725-734.
Leong, S. A., Farman, M. L., Chauhan, R. S., Punekar, N., Mayama, H., Nakayashi, H., Eto, Y., Tosa, Y., and Ronald, P. 2000. Progress on understanding the molecular basis of cultivar specificity in the interaction of the *Magnaporthe grisea AVR1-CO39* with rice variety CO39. Pages 308-315 in: Proc. 2nd Int. Conf. Rice Blast Dis. Kluwer, Dordrecht. The Netherlands.
Leung, H., and Taga, M. 1988 *Magnaporthe grisea* (*Pyricularia* species), the blast fungus. Pages 175-188 in: Advances in Plant Pathology. D. S. Ingram and P. H. Williams, eds. Academic Press, London.
Leung, H., Borromeo, E. S., Bernado, M. A., and Notteghem, J. L. 1988. Genetic analysis of virulence in the rice blast fungus *Magnaporthe grisea*. Phytopathology 78:1227-1233.
Levy, M., Romao, J., Marchetti, M. A., and Hamer, J. E. 1991. DNA fingerprinting with a dispersed repeated sequence resolves pathotype diversity in the rice blast fungus. Plant Cell 3:95-102.
Mackill, D. J., and Bonman, J. M. 1992. Inheritance of blast resistance in near-isogenic lines of rice. Phytopathology 82:746.
Mandel, M. A., Crouch, V. W., Gunawardena, U. P., Harper, T. M., and Orbach, M. J. 1997. Physical mapping of the *Magnaporthe grisea AVR1-MARA* locus reveals the virulent allele contains two deletions. Mol. Plant-Microbe Interact. 10:1102-1105.
Nakayashiki, H., Kiyotomi, K., Tosa, Y., and Mayama, S. 1999. Transposition of the retrotransposon MAGGY in heterologous species of filamentous fungi. Genetics 153:693-703.
Noma, K., Nakajima, R., Ohtsubo, H., and Ohtsubo, E. 1997. RIRE1: A retrotransposon from wild rice *Oryza australiensis.* Genes Genet. Syst. 72:131-140.
Orbach, M. J., Farrall, L., Sweigard, J. A., Chumley, F. G., and Valent, B. 2000. A telomeric avirulence gene determines efficacy for the rice blast resistance gene *Pi-ta.* Plant Cell 12:2019-2032.
Rohe, M., Gierlich, G., Herrmann, H., Hahn, M., Schmidt, B., Rosahl, S., and Knogge, W. 1995. The race specific elicitor NIP1, from the barley pathogen *Rhynchosporium secalis*, determines avirulence on host plants of the *Rsr1* resistance genotype. EMBO (Eur. Mol. Biol. Organ.) J. 14:4168-4177.
Roumen, E. C. 1992. Small differential interactions for partial resistance in rice cultivars to virulent isolates of the blast pathogen. Euphytica 64:143-148.
Smith, J. R., and Leong, S. A. 1994. Mapping of a *Magnaporthe grisea* locus affecting rice (*Oryza sativa*) cultivar specificity. Theor. Appl. Genet. 88:901-908.
Thornbury, D. W., and Farman, M. L. 2000. Reuse of nylon membranes for radioactive hybridizations. Biotechniques 29:1250-1254.
Valent, B., Farrall, L., and Chumley, F. G. 1986. Genetic studies of fertility and pathogenicity in *Magnaporthe grisea* (*Pyricularia oryzae*). IA St. J. Res. 60:569-594.
Valent, B., Farrall, L., and Chumley, F. G. 1991. *Magnaporthe grisea* genes for pathogenicity and virulence identified through a series of backcrosses. Genetics 127:87-101.
Van den Ackerveken, G. F. J. M., Van Kan, J. A. L., and De Wit, P. J. G. M. 1992. Molecular analysis of the avirulence gene *avr9* of the fungal tomato pathogen *Cladosporium fulvum* fully supports the gene for gene hypothesis. Plant J. 2:359-366.
Wang, G.-L., Mackill, D. J., Bonman, M., McCouch, S. R., Champoux, M. C., and Nelson, R. J. 1994. RFLP mapping of genes conferring complete and partial resistance to blast in a durably resistant rice cultivar. Genetics 136:1421-1434.
Zeigler, R. S., Cuoc, L. X., Scott, R. P., Bernado, M. A., Chen, D. H., Valent, B., and Nelson, R. J. 1995. The relationship between lineage and virulence in *Pyricularia grisea* in the Philippines. Phytopathology 85:443-451.

AUTHOR-RECOMMENDED INTERNET RESOURCES

The North Carolina State University *Magnaporthe grisea* BAC-end database: http://152.14.13.200/index.html.