Chromosome Walking to the *AVR1-CO39* Avirulence Gene of *Magnaporthe grisea*: Discrepancy Between the Physical and Genetic Maps

Mark L. Farman*,[1] and Sally A. Leong*,†

**Department of Plant Pathology and †USDA-ARS Plant Disease Resistance Research Unit, University of Wisconsin, Madison, Wisconsin 53706*

Manuscript received September 9, 1997
Accepted for publication July 2, 1998

ABSTRACT

The *avrCO39* gene conferring avirulence toward rice cultivar CO39 was previously mapped to chromosome *1* of *Magnaporthe grisea* between cosegregating markers *CH5-120H* and *1.2H* and marker *5-10-F*. In the present study, this region of the chromosome was physically mapped using RecA-mediated Achilles' cleavage. Cleavage of genomic DNA sequences within *CH5-120H* and *5-10-F* liberated a 610-kb restriction fragment, representing the physical distance between these markers. Chromosome walking was initiated from both markers but was curtailed due to the presence of repetitive DNA sequences and the absence of overlapping clones in cosmid libraries representing several genome equivalents. These obstacles were overcome by directly subcloning the target region after release by Achilles' cleavage and a contig spanning *avrCO39* was thus assembled. Transformation of two cosmids into a virulent recipient strain conferred a cultivar-specific avirulence phenotype thus confirming the cloning of *avrCO39*. Meiotic crossover points were unevenly distributed across this chromosomal region and were clustered around the *avrCO39* locus. A 14-fold variation in the relationship between genetic and physical distance was measured over the *avrCO39* chromosomal region. Thus the poor correlation of physical to genetic distance previously observed in *M. grisea* appears to be manifested over relatively short distances.

MAGNAPORTHE grisea causes a devastating disease of rice known as rice blast. Whether the fungus is able to grow on a rice cultivar is determined by the interaction of avirulence gene (*AVR*/*avr*) products in the fungus with resistance gene products in the host (Keen 1990). Fungal strains possessing an avirulence gene(s) are unable to grow on rice cultivars containing a corresponding resistance gene(s). Typically, resistant rice cultivars do not withstand more than one or two years of cultivation without succumbing to blast (Bonman 1992), suggesting that avirulence is rapidly lost in *M. grisea* populations or that an already virulent component of the population becomes selected. Studying the molecular genetic basis of avirulence and specifically how avirulence is lost will likely illuminate processes involved in the generation and spread of virulent forms of the fungus.

Smith and Leong (1994) reported the mapping of the *avrCO39* gene that causes *M. grisea* strain 2539 to be avirulent on rice cultivar CO39. We describe here the physical mapping and cloning of this gene using a chromosome walking strategy that employed the RecA-mediated Achilles' cleavage (RecA-AC) technique (Ferrin and Camerini-Otero 1991; Koob *et al.* 1992). RecA-AC enables the genome to be cleaved uniquely at predetermined sites by using a RecA protein:oligonucleotide complex to protect chosen restriction sites from methylation. These sites are subsequently cleaved by the corresponding restriction enzyme. By transforming a virulent strain to be avirulent, we have shown that the *avrCO39* gene confers avirulence in a dominant manner. This observation and the relationship of the cloned gene to other previously identified avirulence genes led us to redesignate the gene *AVR1-CO39* to reflect these properties.

Chromosome walking has been instrumental in the cloning of genes from many organisms. For example, several human disease genes have been cloned using such approaches. Until recently, it was not possible to accurately assess the physical distance between markers flanking the target gene, and chromosome walks were initiated without prior knowledge of the distance to be covered. In well-developed systems for which YAC libraries are available, this may not be a serious concern because large sections of the genome can be covered in each step. However, if cosmid and λ libraries are to be used, it is important to have knowledge of the distance involved, as it may be more expeditious to identify closer markers.

Several chromosome walks have been performed in filamentous fungal genomes such as those of *Neurospora crassa* (Kang and Metzenberg 1990; Davis *et al.* 1994; Mautino *et al.* 1993; Rosa *et al.* 1997), *Aspergillus nidulans* (Hull *et al.* 1989), and *Schizophyllum commune* (Gias-

Corresponding author: S. A. Leong, Department of Plant Pathology and USDA-ARS Plant Disease Resistance Research Unit, University of Wisconsin, Madison, WI 53706. E-mail: sal@plantpath.wisc.edu

[1] *Present address:* Department of Plant Pathology, University of Kentucky, Lexington, KY 40546.

son *et al.* 1989). From these studies, some information has been gathered regarding the relationship of genetic to physical distance in filamentous fungal genomes. This information might seem like a useful reference when one is preparing to use chromosome walking to clone genes from other filamentous fungi. However, we show here that this information is of limited practical value in inferring distances between a flanking marker and a target gene. We previously demonstrated that the relationship of genetic to physical distance within the *M. grisea* genome can vary by at least threefold when measured over large chromosomal regions (Farman and Leong 1995). In the present study this relationship was examined over a 610-kb contiguous stretch of *M. grisea* chromosome *1* by mapping meiotic recombination points encountered in a chromosome walk that resulted in the cloning of the *AVR1-CO39* avirulence gene.

MATERIALS AND METHODS

Bacterial and fungal strains and plasmids: *Escherichia coli* strain DH5α (GIBCO BRL, Gaithersburg, MD) was used for all routine cloning experiments. *M. grisea* strains Guy11 and 2539 and 61 of their ascospore progeny have been described previously (Skinner *et al.* 1993). An additional 48 (non-*buf1*$^-$) progeny isolates were generated by M. L. Farman (Smith and Leong 1994).

Design of synthetic oligonucleotides for Achilles' cleavage experiments: Marker *CH5-120H* contained two *Eco*RI sites. DNA sequence surrounding one of these sites was determined by subcloning the insert to place the site near a universal priming site in pBluescript KSII$^+$ (Stratagene, La Jolla, CA).

All other markers were specified by cosmid clones. In these cases the cosmids were subjected to restriction analysis to identify *Bam*HI fragments containing at least one *Eco*RI site. Candidate *Bam*HI fragments were arbitrarily subcloned. The subclones were then restriction mapped to identify restriction sites within 300 bp of the *Eco*RI site in the subclone. These sites were then employed to generate further subclones to enable the DNA sequence spanning the *Eco*RI site to be determined using T3 and T7 primers flanking the multiple cloning site. Oligomers used were as follows: *CH5-120H*: 5′ GGCGGGG GTCCAGAACGCTGATGTTCTCGTCGTTGGCATAGTCGAC CGATGATCTCTTGGAATTCCGGTCGG 3′; *5-10-F*: 5′ CGAA TCCTCGTCGGCATTACCACTGGCAGGTTGGATGACGAG GAGGTTCTGGAATTCGCAGG 3′; *43-2-H*: 5′AATCATTACT CTCATCACTCACCATACGCTTCGGCCTACACATCACAT CCGAATTCCGCATC 3′; and *18-2-F*: 5′ TGACGGTTGATTTA TCCGACCCTGCATCTCAGTTCAGGTGTGCAAGCCTGGA ATTCTGGCGT 3′.

RecA-assisted Achilles' cleavage: Reactions were performed on 100 μl of microbead-embedded DNA (~2 μg), prepared by suspending ~5×10^8 protoplasts per milliliter of molten InCert agarose (FMC Bioproducts, Rockland, ME) and vortexing the suspension in mineral oil as described by Koob *et al.* (1992). The RecA-AC protocol was modified slightly to accommodate the use of two oligonucleotides: The nucleoprotein complex was formed by mixing 2 μl of each oligonucleotide (0.165 μg/μl) to 4 μl of RecA protein (~6 μg/μl; New England Biolabs, Beverly, MA) in 20 μl of the standard 1× reaction buffer. The whole reaction mixture was then added to the microbeads and diffusion and subsequent methylation were performed as described previously (Koob *et al.* 1992). Chromosomal fragments liberated by RecA-AC were resolved from the remaining chromosomal DNA by contour-clamped homogeneous electric field (CHEF) electrophoresis. After photographing, the CHEF gels were depurinated, denatured, and neutralized according to standard protocols (Sambrook *et al.* 1989) and transferred to Magnagraph membranes (MSI). Probes used for hybridization to resolved fragments were previously mapped RFLP marker DNAs (*1.2H*, *CH5-131H*, *pTEL1.8B*) or cosmids identified in this study.

Chromosome walking strategy: An ordered genomic DNA library of *M. grisea* strain 2539, consisting of 5184 clones, was constructed in cosmid vector pMLF1 (Leong *et al.* 1994). This library was stored in microtiter plates and DNA preparations were made by pooling colonies from each microtiter plate. The DNA from individual clones in each plate was also gridded in an 8 × 12 array on nylon membranes by colony blotting (Sambrook *et al.* 1989). Library screening was performed using a modification of a protocol by Bowden *et al.* (1988): Southern blots of the digested DNA pools were probed first to identify microtiter plates containing clones of interest. The individual clones were then identified by a second-round hybridization of the appropriate colony blot. A second library representing approximately 1000 genome equivalents was constructed in pMLF2 (An *et al.* 1996). A second ordered library of 1728 clones was prepared in microtiter plates and the remainder was stored as an unamplified phage suspension. Rare clones were identified by plating and screening aliquots of the unamplified library. Chromosome walking was performed using endclones prepared from the insert DNA by digesting the cosmid clones with *Apa*I, which does not digest the vector, and recircularizing the plasmid by ligation. This procedure results in a derivative containing DNA from each end of the insert (An *et al.* 1996). Liberation of both ends of the insert from the vector was achieved by digesting with *Apa*I and *Not*I. The required endclone was then identified through its failure to hybridize with the previous cosmid in the walk.

Preparation of subclones of a restriction fragment released by Achilles' cleavage: Following CHEF electrophoresis, the 310-kb Achilles' cleavage product was briefly visualized by long wavelength UV illumination of the ethidium bromide-stained gel. A gel slice (~50 μl) containing the fragment was excised and equilibrated in 1 ml of TE buffer for 1 hr. The gel slice was then washed in 1 ml of restriction enzyme buffer (New England Biolabs buffer 3) for a period of 1 hr. This step was repeated for a total of three washes. Finally, excess buffer was removed, leaving just enough to cover the gel slice. Twenty units of *Bam*HI (New England Biolabs) was then added and digestion was performed overnight at 37°. After digestion, the gel slice was equilibrated with TAE buffer. The DNA was then extracted from the gel slice using the Gene Clean procedure (BIO 101, Vista, CA).

Ten microliters of the purified restriction fragments was ligated overnight at 16° in a reaction volume of 20 μl with pBluescript KSII$^+$ (Stratagene) that had been treated with *Bam*HI and calf intestinal alkaline phosphatase (New England Biolabs). Ten microliters of this ligation mixture was transformed into DH5α (GIBCO BRL) using a standard protocol (Sambrook *et al.* 1989).

RFLP mapping of cosmids containing subclones of the 290-kb Achilles' cleavage fragment: Clones used to identify corresponding cosmids from the library had been shown to originate from the target region of the *M. grisea* genome by hybridization to Southern blots of CHEF-resolved cleavage products (Figure 4). These cosmids were mapped genetically within the *AVR1-CO39* locus by determining RFLP marker segregation in 11 progeny that were recombinant in the genetic interval containing *AVR1-CO39*. Although the remaining 50 nonrecombinant progeny were not informative for mapping cosmids within the *CH5-120H* to *5-10-F* interval, 7 of these progeny

were included to reduce the likelihood of being misled by unexpected occurrences such as jumping to other chromosomes through repetitive DNA or chimeric cosmids.

Transformation of virulent strain Guy11 with cosmids within the *AVR1-CO39* locus: Cosmids from within the genetic interval containing *AVR1-CO39* were introduced into Guy11 using the transformation protocol described in Leung *et al.* (1990). The procedure was modified as follows: After the protoplasts were incubated in complete medium (CM) + sorbitol, they were poured into 100 ml molten (45°) CM + 20% sucrose agar. The agar was then poured into four petri plates. When the agar had solidified (1 hr) it was overlaid with 15 ml of 1.5% water agar containing 800 μg/ml hygromycin B (300 μg/ml final concentration).

Physical mapping of meiotic crossovers: At each step of the walk, whole cosmids were labeled by nick translation and used as probes to screen for RFLPs between parental DNAs cut with five restriction enzymes (*Bam*HI, *Dra*I, *Eco*RI, *Hin*dIII, and *Pst*I). Informative probe/enzyme combinations were then used to survey RFLP inheritance in recombinant progeny. For each recombinant progeny isolate, crossovers were revealed when adjacent cosmids identified RFLPs from opposite parents or when a single probe revealed RFLPs from each parent. These crossovers were localized to specific regions by assuming that they occurred midway between polymorphic restriction fragments, whose locations were based on several criteria: (i) If a RFLP was not shared by overlapping cosmids, it was concluded that the polymorphic fragment lay within the central portion of the cosmid insert. Conversely, if a RFLP was shared, it lay within the overlapping portion. (ii) If, in a single progeny DNA sample, a subset of the total RFLPs was inherited from each parent, it was concluded that the crossover point occurred within the region spanned by the cosmid insert. The locations of the informative RFLPs were again determined on the basis of whether they were shared by overlapping cosmids. For most crossovers, the level of resolution thus achieved was to within a window of ~20 kb. For the sliding window analysis, the locations of crossovers were assumed to lie at the midpoint of the window.

Analysis of the distribution of crossovers: A graphical representation of crossover distribution across the 610-kb chromosome segment was obtained by plotting numbers of crossovers occurring in a 50-kb window that was slid in 20-kb steps across the region under study. The relationship of physical to genetic distance across the *CH5-120H* to *5-10-F* interval was calculated as the distance between these markers (in kilobases) divided by the genetic distance (in centimorgans). The genetic distance was calculated by applying the Kosambi mapping function (Kosambi 1944) to the recombination fraction (no. crossovers/total progeny). Assessment of variation in the relationship between genetic and physical distances across a chromosome segment is subject to experimental bias in the choice of size and position of the intervals across which these values are measured. To minimize this bias, a series of sliding window analyses was performed using three different window widths and the ratios of physical to genetic distance were compared for window positions across the entire chromosome region under study. The physical distance in these analyses equaled the window width and the genetic distance at each window position was calculated on the basis of the number of crossovers occurring in the window.

RESULTS

Physical mapping of the end of the chromosome containing *AVR1-CO39*: Achilles' cleavage of the 2539 genome resulted in the liberation of segments of the chromosome between the targeted cleavage sites. These fragments were resolved from the uncut chromosomes by CHEF electrophoresis. The conditions chosen for optimal separation of the cleavage products did not allow resolution of the uncut chromosomes that migrated as a single band in these experiments. In some cases, the cleavage products did not show up clearly in the agarose gel but were readily identified when Southern blots were hybridized with appropriate probes (Figure 1).

Achilles' cleavage within marker *CH5-120H* yielded a restriction fragment that was clearly resolved from the remaining intact chromosomes by CHEF electrophoresis and Southern hybridization with appropriate probes (Figure 1). The size of this fragment was estimated as 1.29 mb by virtue of its comigration with the correspondingly sized chromosome of *Hansenula wingeii* (Figure 1A and B). Marker *1.2H* (Skinner *et al.* 1993) and telomere-associated sequence, *TEL1.8B* (Farman and Leong 1995) also hybridized to this restriction fragment (Figure 1B and C). These results confirmed that this fragment extends to the telomere and indicated that marker *1.2H* lies distal to *CH5-120H* and hence closer to *AVR1-CO39*. Simultaneous cleavage of the 2539 genome at markers *CH5-120H* and *5-10-F* yielded three restriction fragments that were resolved from the intact chromosomes. The 1.29-mb fragment which hybridized to the *1.2H* and *TEL1.8B* probes was generated by incomplete protection from methylation of the *Eco*RI site in *5-10-F*. The second fragment of ~680-kb hybridized only to the pTEL1.8B probe and contains the telomere indicating that marker *5-10-F* lies 680 kb from the end of the chromosome. The final fragment, which was 610 kb in size, hybridized to marker *1.2H* (Figure 1B) and to one endclone of *5-10-F* (results not shown). This confirmed that this fragment also contained *AVR1-CO39*. Marker *CH5-131H*, which maps telomere distal to CH5-120H (Skinner *et al.* 1993), was included as a control probe. As expected, this probe hybridized only to the unresolved chromosome band (Figure 1D), which is consistent with hybridization to the remainder of chromosome *1*. As a result of these experiments, we were able to physically map *AVR1-CO39* to a 610-kb region, which was delimited by markers *CH5-120H* and *5-10-F*.

Chromosome walking towards *AVR1-CO39*: A bidirectional chromosome walk was initiated from marker *1.2H* in both directions. Cosmid *21-9-D*, containing *1.2H*, overlapped with *16-1-F*, which identified a RFLP that mapped one map unit closer to *AVR1-CO39*. This recombination event established the correct orientation of the walk. The direction of the walk from marker *5-10-F* was rapidly established by determining which of the two endclones from this cosmid hybridized to the 610-kb Achilles' cleavage product (results not shown). Nine steps from marker *1.2H*, a cosmid was identified that contained a portion of the *GRASSHOPPER* retroelement (Dobinson *et al.* 1993) at the distal end of the insert (Figure 2A). This element is present in more than 30

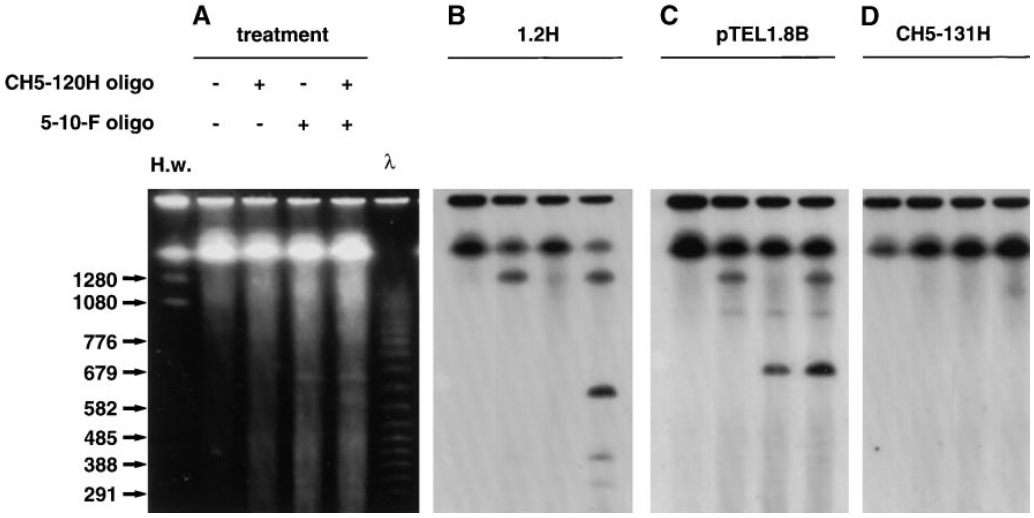

Figure 1.—Physical mapping of markers flanking *AVR1-CO39*. Achilles' cleavage was performed using oligonucleotides specific for sequences in markers *CH5-120H* and *5-10-F*. Single-site cleavages were made at markers *CH5-120H* and *5-10-F* to determine the distances from these markers to the telomere. Simultaneous cleavage within both markers was employed to determine the distance between them. Achilles' cleavage reactions were performed as described in materials and methods. The products were electrophoresed in a 1% FastLane agarose gel (0.5× TBE) using the following CHEF conditions: 150V, 60 sec switch time, 16 hr, followed by 140 sec switch, 4 hr. The voltage was increased to 200 V, the switch time was ramped from 60 sec to 140 sec for 6 hr and finally a 140 sec switch time was applied for 3 hr. (A) A photograph of the gel. (B–D) The gel was blotted to a nylon membrane and hybridized sequentially with the probes indicated above each autoradiogram. H.w., *H. wingei* chromosomal size standard; λ, a lambda concatemer. Molecular sizes are indicated in kilobases.

copies in the 2539 genome (M. Farman and S. A. Leong, unpublished observations) and the endclone identified more than 50 clones in the 2539 library. A DNA fragment proximal to *GRASSHOPPER* was used to identify the clone *cos3-7* that extended beyond the element. However, these sequences were also present in multiple copies in the 2539 genome. Thus the chromosome walk from *1.2H* was curtailed at this point resulting in a gap in the contig shown as a gray box in Figure 2A.

The chromosome walk from *5-10-F* extended five steps and spanned four recombination points in the initial progeny population (Figure 2A). However it was not possible to identify a clone overlapping 18-2-F in the 2539 DNA library. A second library was constructed in pMLF2 and screened to bridge the gap, but no overlapping clones were identified in more than 10,000 additional cosmids surveyed.

Subcloning a restriction fragment released by Achilles' cleavage of chromosome *1* at markers *43-2-H* and *18-2-F*: Oligonucleotide primers were designed to overlap restriction sites in the cosmids *43-2-H* and *18-2-F*, which were identified while chromosome walking towards *AVR1-CO39* (indicated by triangles in Figure 2B). Achilles' cleavage of the chromosome at these sites resulted in the liberation of a 310-kb *Eco*RI fragment. This fragment was excised, digested with *Bam*HI, and subcloned into pBluescript KSII$^+$. A total of 104 subclones was obtained but it was anticipated that a proportion of the clones obtained in this manner would not be derived from the Achilles' cleavage product because some shearing had occurred during preparation of the chromosomal DNA (see Figure 1). Therefore, subclones were used as probes to Southern blots of CHEF gels in which the 310-kb fragment was resolved. Out of the first 12 subclones tested, 5 (subclones 3, 4, 7, 11, and 16) were derived from the target region and were judged to be single copy on the basis of their relative hybridization intensity to the 310-kb fragment and the uncut chromosome band (results not shown). The inserts in these clones were then used as probes to identify corresponding cosmids in the ordered library constructed in pMLF2.

Map locations of cosmids identified by subclones of the 310-kb Achilles' cleavage product: RFLP analysis using entire cosmid clones as probes indicated that there was a considerable degree of polymorphism between the parental strains in this region of chromosome *1*. A cosmid that contained subclones 3, 7, and 11 produced similar RFLP profiles to cosmid *16-5-1* and mapped at the same locus. Subclone 4 identified two cosmids, *7-4-D* and *17-4-B*, the latter of which identified four RFLPs (Figure 3A). Most progeny inherited all RFLPs from one or the other parent but isolates 6050, 6068, and 6081 inherited various subsets of these RFLPs from each parent. This indicated that the meiotic recombination events had actually occurred within the chromosomal region encompassed by the cosmid insert. As a result, one of the RFLPs (#1) cosegregated fully with *AVR1-CO39* in the original mapping population and occurred 0.9 cM away when the progeny population was expanded for studying crossover distribution. The other RFLPs identified by *17-4-B* map 1.8, 2.7, and 3.6 cM closer to *5-10-F* (Figure 3B). Several cosmids were identified that contained subclone #16. One of these, *cos6B*, identified three RFLPs. One RFLP again cosegregated fully with *AVR1-CO39* in the original progeny population (0.9 cM away in the expanded progeny population) while the others mapped 1.8 and 3.6 cM closer to *1.2H*. We con-

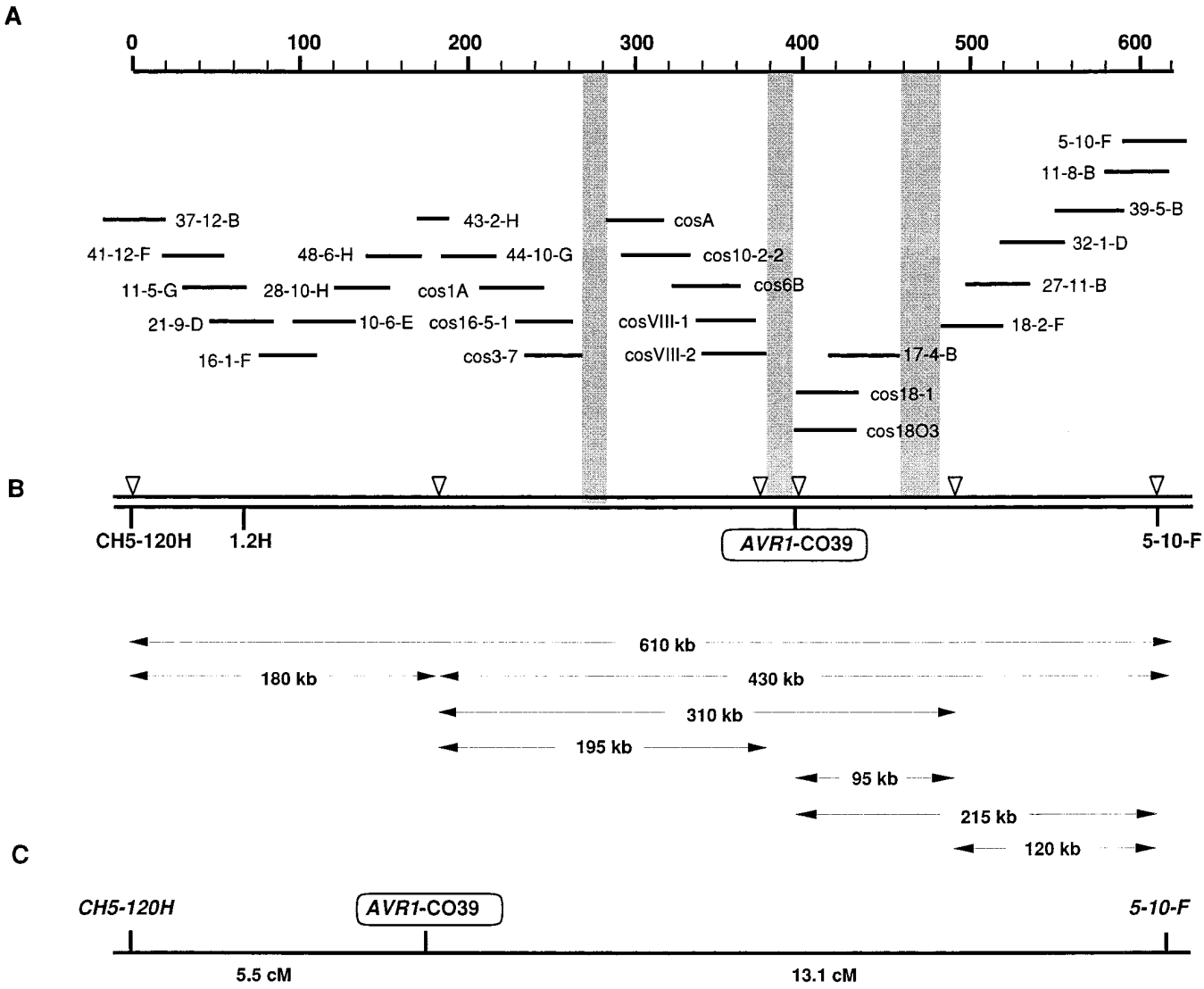

Figure 2.—Genetic and physical maps of the 610-kb chromosomal region encompassing *AVR1-CO39*. (A) Cosmid contigs assembled while walking to *AVR1-CO39*. Representative cosmids with minimal overlap are shown. Gray shaded areas represent gaps in the chromosome walk due to repetitive or "uncloneable" DNA. (B) Summary of the physical map of the region as determined by Achilles' cleavage reactions. Cleavage sites are denoted by open arrowheads and the sizes of fragments released are shown. (C) Genetic map of the *AVR1-CO39* locus between markers *CH5-120H* and *5-10-F*. Shown are revised map distances (see text for explanation) which are based on segregation of RFLPs obtained with the cloned *AVR1-CO39* gene probe. Map distances are shown in centimorgans (Kosambi mapping function).

cluded from these studies that cosmids *17-4-B* and *cos6B* flank *AVR1-CO39* on the 5-10-F and 1.2H proximal sides, respectively.

Hybridization analysis established that there was not yet an overlap between *cos6B* and *17-4-B*. Therefore, conventional chromosome walking was resumed to try to establish a contig encompassing the *AVR1-CO39* gene. No cosmids extending beyond *17-4-B* were identified among a total of 6912 clones in two ordered libraries. A total of 20,000 additional colonies (~20 genome equivalents) were screened and a single cosmid, *cos18-1*, was identified that extended ~16 kb beyond *17-4-B*. The endclone of *cos18-1* identified 13 cosmids in over 20,000 additional colonies screened and, although the ends of the inserts in many of these cosmids lay within 5 kb of the end of the contig, only one of these, *cos18O3*, extended further than *cos18-1* and only by 800 bp. While walking towards this region from *cos6B*, several overlapping cosmids were identified. However, those that extended any significant distance beyond *cos6B* (including VIII-2) appeared to have been rearranged as their restriction maps were incongruent with a long-range physical map of the region (results not shown). Nevertheless, the end of the insert in an intact clone, *cosVIII-1*, was ~3 kb closer to *AVR1-CO39*. Together, these observations suggest that the chromosome region beyond *cos18O3* and *cosVIII-1* is not clonable in an intact form using a cosmid vector.

In total, over 100 independent cosmids were isolated

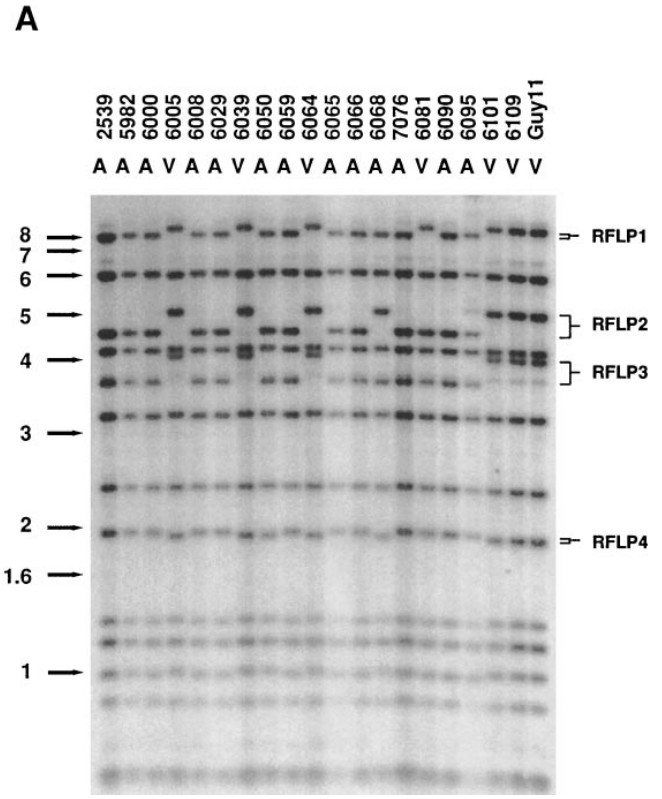

Figure 3.—Segregation of RFLPs identified by cosmid *17-4-B*. (A) *Pst*I-digested DNA of each parental isolate was loaded in the outside lanes as indicated and DNAs of 18 representative progeny (including 11 of 12 that were recombinant in the interval between markers *CH5-120H* and *5-10-F*) were loaded in the middle. RFLPs used to map *17-4-B* are indicated. The infection phenotypes of each isolate are indicated: A, avirulent; V, virulent. (B) Map locations of cosmids in the chromosome walk to *AVR1-C039*. Cosmids *cos6B* and *17-4-B* each identified several polymorphic fragments that were mapped individually and each RFLP is denoted by a suffix after the marker name. The map was created using MAPMAKER v2.0 (Lander *et al.* 1987) and the Kosambi mapping function was used.

during the walking process and over 500 kb of the chromosome was covered. A summary of the walk is shown in Figure 2A in which minimally overlapping clones and other landmark cosmids are depicted.

Transformation of Guy11 to avirulence with cosmids containing *AVR1-CO39*: The 4.3-kb *Bam*HI to *Not*I endclone of *cos18O3* did not hybridize to genomic DNA of the virulent *M. grisea* strain Guy11 (results not shown), suggesting that part of the *AVR1-CO39* locus was deleted in this strain. It was hypothesized that Guy11 may have gained virulence to CO39 through deletion of the *AVR1-CO39* gene. Therefore, protoplasts of Guy11 were transformed with cosmids possessing DNA that was within the deletion. Cosmids 18O3 and 18-1 both converted Guy11 to avirulence and transformants carrying these clones were unable to infect CO39. Their capability to infect cultivar 51583 was completely unaffected, confirming that the gene acted in a cultivar-specific manner. Transformants carrying an *Apa*I deletion derivative of 18O3, "18O3ΔA," were also able to confer avirulence to CO39 (Figure 4) and localized *AVR1-CO39* activity to a 7.15-kb *Apa*I to *Not*I subclone of *cos18O3*. Transformants carrying cosmid *cos18OA*, which bears *AVR1-CO39* with a deletion in a presumed promoter region, produced intermediate-sized lesions (results not shown), while transformants harboring cosmids *cosVIII-2*, *cos11-1*, and *cos3-2*, all of which lack *AVR1-CO39* sequences, were unaffected in their infectivity to either CO39 (data not shown) or 51583.

Physical mapping of cosmid contigs: As a result of directly subcloning portions of the chromosome from within the target region, it was not known how the new cosmid contigs centered around *cos6B* and *17-4-B* were physically disposed with respect to those assembled in the initial chromosome walk. Therefore, it was not possible to physically map recombination points that were identified by *cos6B*, *cosVIII-1*, *cos18O3*, and *17-4-B*. This limitation was overcome by applying Achilles' cleavage to map the distance between *Eco*RI sites within: (i) *cosVIII-1* and *43-2-H* and (ii) *cos18O3* and *18-2-F*. The sizes of the respective cleavage products revealed that *cos VIII-1* lies 195 kb from *43-2-H* and *cos18O3* is 95 kb from *18-2-F*. Additional cleavage reactions were performed with different oligonucleotide combinations to confirm these distances. A summary of these results and the physical map derived by the chromosome walking and Achilles' cleavage studies is shown in Figure 2B. By combining these data with distances determined by chromosome walking it was possible to create a comprehensive physical map of the *CH5-120H* to *5-10-F* region (Figure 2B). These analyses revealed that *AVR1-CO39* was considerably closer to *5-10-F* than was expected based on its map location (Figure 2C).

Revised map location of *AVR1-CO39*: The published map location of *AVR1-CO39* is 11.8 cM from *CH5-120H* and 17.2 cM from *5-10-F* (Smith and Leong 1994). These distances were previously approximated because one-quarter (16/61) of the progeny of the Guy11 × 2539 cross exhibited a pigmentation deficiency caused by meiosis-induced deletion of the *BUF1* gene (M. Farman,

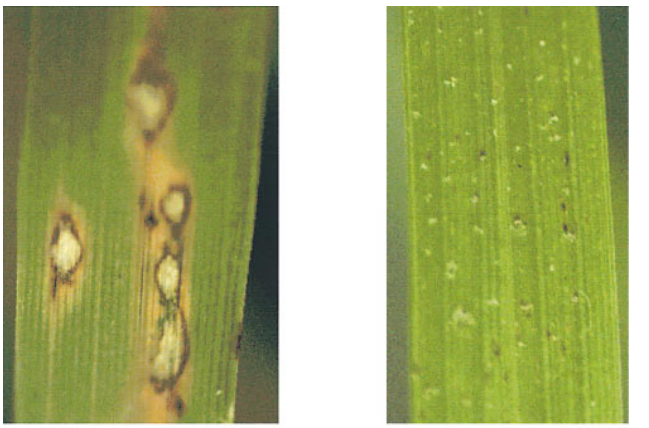

Figure 4.—Infection phenotypes of Guy11 and transformants carrying *AVR1-CO39*. Representative infections obtained with the original recipient strain, an avirulent transformant (18O3ΔA#6), are shown. Conidia of Guy11 and single spore isolates of Guy11 transformants, 18O3ΔA#6, were sprayed onto seedlings of rice cultivar CO39 at an inoculum density of 5×10^4 conidia/ml. Prior to inoculation, at least 12 single spore isolates were obtained from each spore sample and were transferred to oatmeal agar containing hygromycin B to confirm the stability of the transforming DNA.

unpublished results). Such mutants are nonpathogenic because they are unable to produce melanin, which appears to be an essential component of penetration structures (Bourett and Howard 1990; Chumley and Valent 1990). Consequently, we were unable to evaluate reliably their infection phenotype on CO39. The cloned *AVR1-CO39* gene from the present study provided a molecular marker enabling accurate genotypic assignments to be made for the *buf1*$^-$ progeny. In a previous study to confirm single gene inheritance of avirulence to CO39, 53 progeny were generated in addition to those used for mapping (Smith and Leong 1994). In the course of this study, marker inheritance surrounding *AVR1-CO39* was examined in 48 of these additional progeny. The map location of the *AVR1-CO39* gene was adjusted to account for these new data.

A 4.3-kb *Bam*HI to *Not*I fragment containing the *AVR1-CO39* gene detected three *Dra*I fragments in 2539 but did not hybridize to genomic DNA of Guy11 (results not shown). Segregation analysis of this RFLP among a combined population of 109 progeny indicated that *AVR1-CO39* actually lies 5.5 cM from *CH5-120H* (6 recombinant progeny out of 109) and 13.1 cM from *5-10-F* (14 out of 109). Thus the total genetic distance between these markers is 18.6 cM, which equates to an average physical distance of ~33 kb/cM, or one crossover event every 30.5 kb.

Locations and distribution of crossovers: All exchanges, except one, could be localized within sequences encompassed by cosmid clones. The exceptional crossover event occurred somewhere within the region of highly repetitive DNA and was poorly resolved to a 40-kb chromosome segment. In a related study, an anonymous *GRASSHOPPER* element was mapped to a location beyond this recombination point. For the sliding window analysis, a reasonable assumption was made that the mapped *GRASSHOPPER* corresponds to the element in *cos16-5-1*. It was clear that recombination was not distributed evenly across the *CH5-120H* to *5-10-F* interval. For example, markers *CH5-120H* and *1.2H* cosegregated fully and are ~80 kb apart (1 cM > 80 kb). Similarly, no recombination points were encountered between markers *16-1-F* and *16-5-1*, which are separated by ~150 kb (1 cM > 150 kb). In contrast, the inserts in cosmids *cos6B* and *17-4-B* each detected three crossover points, implying that 1 cM < 10 kb. On the whole, recombination points appeared more evenly distributed within the rightmost cosmid contig and the pattern of recombination across the region was very similar between the first and second crosses (Figure 5A). A graphical representation of this distribution was made by performing a sliding window analysis using a 50-kb window size and plotting the number of crossovers in a window against the window position. This analysis highlighted recombinational clusters surrounding the *AVR1-CO39* locus and at marker *39-5-B* (Figure 5B).

Results obtained from three window sizes (300, 200, and 150 kb) indicated that variation of the window width altered the relationship between genetic and physical distance. This was expected because, at certain positions in the physical map, the increased window width did not encompass additional crossovers. Nevertheless, in all cases, a recombinational gradient was measured across the interval under study with the difference in the kb/cM ratio reaching as high as 14-fold using the 200-kb window in which 1 cM represented as much as 218 kb at the *CH5-120H* end of the interval but equated to 16 kb in regions surrounding the *AVR1-CO39* locus (Figure 5C). Interestingly, the kb/cM ratio was unaffected by window size in the recombinationally active portion of the chromosome region under study (Figure 5C).

DISCUSSION

At least five genetic maps of *M. grisea* have been constructed (Romao and Hamer 1992; Skinner *et al.* 1993; Sweigard *et al.* 1993; Hayashi and Naito 1994; Dioh *et al.* 1996). The last four maps were specifically constructed with the intention of using map-based strategies to clone avirulence genes. As a result at least 10 loci conferring cultivar or host specificity have been mapped (Sweigard *et al.* 1993; Smith and Leong 1994; Hayashi and Naito 1994; Kang *et al.* 1995; Dioh *et al.* 1996). These studies have laid the foundation for the isolation of these genes through systematic map-based cloning

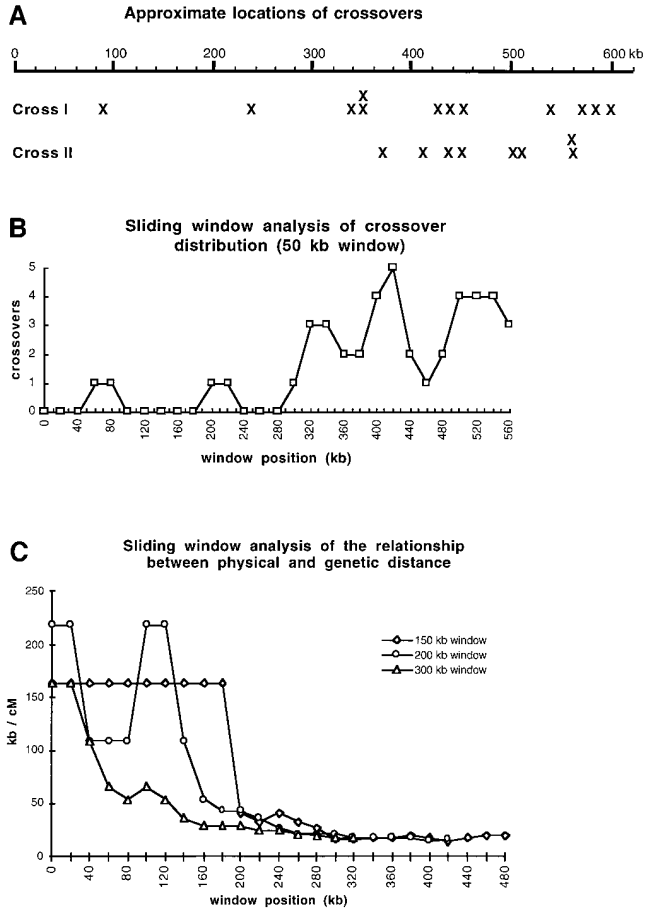

Figure 5.—Distribution and sliding window analysis of recombination across a 610-kb region of the *M. grisea* genome. The distribution of meiotic crossovers is shown in (A). For the sliding window analyses, the approximate locations of meiotic crossover points were determined by assuming they occurred midway between the RFLPs that revealed each recombination event. Each crossover point is marked with an X. (Top) Crossovers occurring in cross 1; (bottom) those occurring in cross 2. (B) Window position was plotted against the number of crossovers occurring in a 50-kb window. The window position, displayed on the x-axis, represented the distance from the left corner of the window to *CH5-120H*. (C) Window position was plotted against the ratio of physical to genetic distance measured across windows representing chromosomal regions of 150, 200, and 300 kb.

approaches. The cultivar specificity genes *AVR2-YAMO* and *PWL2* were both isolated through the knowledge of their chromosomal locations. The *AVR2-YAMO* gene was mapped to a telomere and cloning of the corresponding chromosome end resulted in the isolation of this gene (Valent and Chumley 1994). In contrast, *PWL2* was linked to an internal RFLP marker. A chromosome walk from this marker resulted in the cloning of *PWL2* (Sweigard *et al.* 1995).

The genetic distance of 18.6 cM between the markers flanking *AVR1-CO39* seems large when compared to distances in organisms with larger genomes, where 1 cM can represent 1 Mb. However, earlier physical mapping studies at the other end of chromosome *1* had indicated that 1 cM represents approximately 50 kb (Farman and Leong 1995), a value that is similar to those measured in other filamentous fungi such as *A. nidulans* (3–4 kb/cM, Hull *et al.* 1989; 15 kb/map unit, Lints *et al.* 1995) and *N. crassa* (10–80 kb/map unit, Kang and Metzenberg 1990; Carribo *et al.* 1991; Aronson *et al.* 1992; Mautino *et al.* 1993).

A physical distance of 610 kb between flanking markers *CH5-120H* and *5-10-F* indicated that a chromosome walk of ~300 kb from each marker was certain to result in the cloning of the gene. As this physical distance was not overly large, we chose to initiate a walk rather than try to identify closer markers. At the inception of the walk we were led astray by the genetic proximity of *AVR1-CO39* to markers *CH5-120H* and *1.2H* and as a result, a greater emphasis was placed on assembling a contig from these markers. In retrospect this was an unwise strategy due to the uneven distribution of recombination points, and we hope that our findings will alert others to the pitfalls associated with such an approach.

The cloning of *AVR1-CO39* described here represents one of the most extensive walks performed in *M. grisea* to date, and one of the largest walks in a filamentous fungal genome. All possible obstacles were encountered including occasional chimeric clones, complex arrays of repetitive elements, and chromosome regions that were apparently uncloneable in *E. coli*. Interestingly, Mandel *et al.* (1997) also reported finding uncloneable DNA sequences adjacent to an *M. grisea* avirulence gene. Fortunately, in our case, the ability to cleave the genome using RecA-AC enabled these problems to be rapidly surmounted.

The disparity in the relationship between genetic and physical distance was great. For example, while walking from *1.2H*, over 150 kb was traversed (from cosmids *16-1-F* to *16-5-1*) without encountering a single crossover. By contrast, a crossover was identified in almost every step while walking from *5-10-F*, and cosmids *6B* and *17-4-B*, which defined the boundaries of the *AVR1-CO39* locus, each spanned three crossover points. Over the interval studied, which represents ~1.5% of the 38 Mb *M. grisea* genome (Hamer *et al.* 1989), the ratio of physical to genetic distance was ~1 cM to 33.5 kb. However, regional differences in this ratio varied 14-fold. This variation is greater than that observed in some regions of the *N. crassa* genome proposed to be under the influence of the "centromere effect" (Davis *et al.* 1994) wherein recombination is suppressed in proximity to the centromere. The centromere effect has been borne out by physical mapping studies in *N. crassa* (Rosa *et al.* 1997) where the ratio of physical to genetic distance has been extrapolated to be 22 mb per map unit for regions very close to the centromere (Centola and Carbon 1994). This is a potential cause of the low recombination observed in the present study. However, if a centromere effect caused recombination to be re-

duced over a significant portion of the chromosome, one might expect RFLP markers to cluster in this region of chromosome *1* but this was not the case (Skinner *et al.* 1993; Nitta *et al.* 1997). In contrast, there was a cluster of markers in a region 35 cM from *CH5-120H* (Skinner *et al.* 1993), where eight RFLP probes mapped to a 4.8-cM interval. This part of the chromosome would appear to be a more likely location for the centromere.

Three recombinational clusters were revealed, one on each side of the *AVR1-CO39* locus and one centered around the cosmid marker *39-5-B* (Figure 5C). There was a recombination-deficient area in the middle of the avirulence gene locus that corresponded to a 20-kb deletion/insertion polymorphism between the parents. Clearly this prevents sequences within the deletion/insertion from partaking in crossovers. Interestingly, there appeared to be a compensatory elevation of crossing over in the immediate flanking DNA regions.

The finding that the distribution of crossover points encountered while walking resulted in a poor correlation between the genetic and physical maps is in contrast with the reported distribution of crossovers in other filamentous fungi such as *A. nidulans* and *N. crassa* (Hull *et al.* 1989; Mautino *et al.* 1993). In these fungi there is good correlation between the genetic and physical maps. However, in most of these studies, comparatively short distances were examined and the locations of crossover points were not determined. For example, Mautino *et al.* (1993) mapped 38 crossovers over a 200-kb chromosomal region between the *eth-1* and *un-2* loci of *N. crassa*. Segregation of a total of five RFLP loci was analyzed in their study and crossovers were thus delimited to four intervals. It was assumed that the crossovers were evenly distributed across each interval but it is equally possible that they were clustered between RFLPs. We should qualify this observation by noting that even if clustering had occurred, the distribution of crossovers would not be as polarized as that described in the present study, as recombination was abundant across the entire *eth-1* to *un-2* interval.

The *AVR1-CO39* avirulence gene was first identified by Valent *et al.* (1991) in crosses between the rice-pathogenic isolate O-135 and the weeping lovegrass pathogen 4091-5-8. Strain O-135 was pathogenic on CO39 and therefore they concluded that the 4091-5-8 parent carries *AVR1-CO39*. As 2539 is a descendant of 4091-5-8 (Leung *et al.* 1988), we suspected that the gene we cloned was *AVR1-CO39*. An RFLP analysis of genomic DNA sequences surrounding the *AVR1-CO39* locus in 2539 and 4091-5-8 confirmed this suspicion (results not shown). The numerical designation for the gene is important as Zeigler *et al.* (1995) have identified isolates of *M. grisea* that are avirulent on CO39 yet lack *AVR1-CO39* (M. L. Farman and S. A. Leong, unpublished data). These observations indicate that additional *AVR-CO39* genes exist.

Measurement of disease severity can be prone to subjective interpretation. However, the cloning of *AVR1-CO39* confirmed that the phenotypic scores used to map the gene (Smith and Leong 1994) were accurate. The mapping data predicted that particular progeny should have crossovers between *AVR1-CO39* and flanking markers. All the predicted crossover events were confirmed by the analysis presented herein. The location of the *AVR1-CO39* transcriptional unit is currently being defined through DNA sequencing, deletion analysis, and transcript mapping. Through molecular characterization of *AVR1-CO39* and its functional homologs, we hope to understand the molecular basis of host/cultivar specificity. Furthermore, studies of the evolution of this gene will provide insight into mechanisms affecting the diversification of host range and consequently the breakdown of host resistance.

This work was supported by the United States Department of Agriculture and by a grant from the Rockefeller Foundation to S.A.L.

LITERATURE CITED

An, Z., M. L. Farman, A. D. Budde, S. Taura and S. A. Leong, 1996 New cosmid vectors for library construction, chromosome walking and restriction mapping in filamentous fungi. Gene **176:** 93–96.

Aronson, B. D., K. A. Johnson, Q. Liu and J. C. Dunlap, 1992 Molecular analysis of the *Neurospora clock*: cloning and characterization of the *frequency* and *period* genes. Chronobiol. Int. **9:** 231–239.

Bonman, J. M., 1992 Durable resistance to rice blast disease—environmental influences. Euphytica **63:** 115–123.

Bourett, T. M., and R. J. Howard, 1990 *In vitro* development of penetration structures in the rice blast fungus *Magnaporthe grisea*. Can. J. Bot. **68:** 329–342.

Bowden, D. W., H. Müller-Kahle, T. R. Fulton, T. C. Gravius, D. F. Barker *et al.*, 1988 Studies on locus expansion, library representation, and chromosome walking using an efficient method to screen cosmid libraries. Gene **71:** 391–400.

Carribo, A., E. Sporeno, G. Macino and P. Ballario, 1991 CMB1, a *Neurospora crassa* genomic cosmid library in pAC3 and its use for walking on the right arm of linkage group VII. Fungal Genet. Newsl. **38:** 68–70.

Centola, M., and J. Carbon, 1994 Cloning and characterization of centromeric DNA from *Neurospora crassa*. Mol. Cell. Biol. **14:** 1510–1519.

Chumley, F. G., and B. Valent, 1990 Genetic analysis of melanin-deficient, nonpathogenic mutants of *Magnaporthe grisea*. Mol. Plant-Microbe Interact. **3:** 135–143.

Davis, C. R., R. R. Kempainen, M. S. Srodes and C. R. McClung, 1994 Correlation of the physical and genetic maps of the centrometric region of the right arm of linkage group III of *Neurospora crassa*. Genetics **136:** 1297–1306.

Dioh, W., D. Tharreau, R. Gomez, E. Roumen, M. Orbach *et al.*, 1996 Mapping avirulence genes in the rice blast fungus *Magnaporthe grisea*, pp. 916–920 in *Rice Genetics III*, edited by G. S. Khush. International Rice Research Institute, Philippines.

Dobinson, K. F., R. E. Harris and J. E. Hamer, 1993 *Grasshopper*, a long terminal repeat (LTR) retroelement in the phytopathogenic fungus *Magnaporthe grisea*. Mol. Plant-Microbe Interact. **6:** 114–126.

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

Farman, M. L., Y. Tosa, N. Nitta and S. A. Leong, 1996 MAGGY, a retrotransposon found in the genome of the rice blast fungus, *Magnaporthe grisea*. Mol. Gen. Genet. **251:** 665–674.

Ferrin, L. J., and R. D. Camerini-Otero, 1991 Selective cleavage of human DNA: RecA-Assisted Restriction Endonuclease (RARE) cleavage. Science **254:** 1494–1497.

Giasson, L., C. A. Specht, C. Milgrim, C. P. Novotny and R. C. Ullrich, 1989 Cloning and comparison of *Aa* mating-type alleles of the basidiomycete *Schizophyllum commune*. Mol. Gen. Genet. **218:** 72–77.

Hamer, J. E., L. Farrall, M. J. Orbach, B. Valent and F. G. Chumley, 1989 Host species-specific conservation of a family of repeated DNA sequences in the genome of a fungal plant pathogen. Proc. Natl. Acad. Sci. USA **86:** 9981–9985.

Hayashi, N., and H. Naito, 1994 Genome mapping of *Magnaporthe grisea*, p. 590 in *Rice Blast Disease*, edited by R. S. Zeigler, S. A. Leong and P. S. Teng. CAB International, Wallingford, U.K.

Hull, E. P., P. M. Green, H. N. Arst, Jr. and C. Scazzocchio, 1989 Cloning and physical characterization of the l-proline catabolism cluster of *Aspergillus nidulans*. Mol. Microbiol. **3:** 553–559.

Kang, S., and R. L. Metzenberg, 1990 Molecular analysis of *nuc-1+*, a gene controlling phosphorus acquisition in *Neurospora crassa*. Mol. Cell. Biol. **10:** 5839–5848.

Kang, S., J. A. Sweigard and B. Valent, 1995 The PWL host specificity gene family in the blast fungus *Magnaporthe grisea*. Mol. Plant-Microbe Interact. **8:** 939–948.

Keen, N. T., 1990 Gene-for-gene complementarity in plant-pathogen interactions. Annu. Rev. Genet. **24:** 447–463.

Koob, M., A. Burkiewicz, J. Kur and W. Szybalski, 1992 RecA-AC: Single-site cleavage of plasmids and chromosomes at any predetermined restriction site. Nucleic Acids Res. **20:** 5831–5836.

Kosambi, D. D., 1944 The estimation of map distances from recombination values. Ann. Eugen. **12:** 172–175.

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly *et al.*, 1987 MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomica **1:** 174–181.

Leong, S. A., M. L. Farman, J. A. Smith, A. D. Budde, Y. Tosa *et al.*, 1994 Molecular-genetic approach to the study of cultivar specificity in the rice blast fungus, pp. 87–110 in *Rice Blast Disease*, edited by R. S. Zeigler, S. A. Leong and P. S. Teng. CAB International, Wallingford, UK (in association with the International Rice Research Institute, Philippines).

Leung, H., E. S. Borromeo, M. A. Bernardo and J. L. Notteghem, 1998 Genetic analysis of virulence in the rice blast fungus *Magnaporthe grisea*. Phytopathology **78:** 1227–1233.

Leung, H., U. Lehtinen, R. Karjalainen, D. Skinner, P. Tooley *et al.*, 1990 Transformation of the rice blast fungus, *Magnaporthe grisea* to hygromycin B resistance. Curr. Genet. **17:** 409–411.

Lints, R., M. Davis and M. Hynes, 1995 A chromosome walk linking the *gatA* and alcC genes of *Aspergillus nidulans*. Fungal Genet. Newsl. **42:** 44–45.

Mandel, M. A., V. W. Crouch, U. P. Gunawardena, T. M. Harper and M. J. Orbach, 1997 Physical mapping of the *Magnaporthe grisea AVR1-MARA* locus reveals the virulent allele contains two deletions. Mol. Plant-Microbe Interact. **10:** 1102–1105.

Mautino, M. R., S. D. Haedo and A. L. Rosa, 1993 Physical mapping of meiotic crossover events in a 200 kb region of *Neurospora crassa* linkage group I. Genetics **134:** 1077–1083.

Nitta, N., M. L. Farman and S. A. Leong, 1997 Genome organization of *Magnaporthe grisea*: integration of genetic maps, clustering of transposable elements and identification of genome duplications and rearrangements. Theor. Appl. Genet. **95:** 20–32.

Romao, J., and J. E. Hamer, 1992 Genetic organization of a repeated DNA sequence family in the rice blast fungus. Proc. Natl. Acad. Sci. USA **89:** 5316–5320.

Rosa, A. L., S. D. Haedo, E. D. Temporini, G. A. Borioli and M. R. Mautino, 1997 Mapping chromosome landmarks in the centromere I region of *Neurospora crassa*. Fungal Genet. Cell Biol. **21:** 315–322.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 Molecular Cloning: A Laboratory Manual, Ed. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Skinner, D. Z., A. D. Budde, M. L. Farman, J. R. Smith, H. Leung *et al.*, 1993 Genome organization of the rice blast fungus, *Magnaporthe grisea*: genetic map, electrophoretic karyotype and occurrence of repeated DNAs. Theor. Appl. Genet. **87:** 545–557.

Smith, J. R., and S. A. Leong, 1994 Mapping of a *Magnaporthe grisea* locus affecting rice (*Oryza sativa*) cultivar specificity. Theor. Appl. Genet. **88:** 901–908.

Sweigard, J. A., B. Valent, M. J. Orbach, A. M. Walter, A. Rafalski *et al.*, 1993 Genetic map of the rice blast fungus *Magnaporthe grisea* (n = 7), pp. 3.112–3.114 in *Genetic Maps*, Ed. 6, edited by S. J. O'Brien. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sweigard, J. A., A. Carrol, S. Kang, L. Farrall, F. G. Chumley *et al.*, 1995 Identification, cloning, and characterization of PWL2, a gene for host species specificity in the rice blast fungus. Plant Cell **7:** 1221–1233.

Valent, B., and F. G. Chumley, 1994 Avirulence genes and mechanisms of genetic instability in the rice blast fungus, pp. 111–134 in *Rice Blast Disease*, edited by R. S. Zeigler, S. A. Leong and P. S. Teng. CAB International, Wallingford, UK (in association with the International Rice Research Institute, Philippines).

Valent, B., L. Farrall and F. G. Chumley, 1991 *Magnaporthe grisea* genes for pathogenicity and virulence identified through a series of backcrosses. Genetics **127:** 87–101.

Zeigler, R. S., L. X. Cuoc, R. P. Scott, M. A. Bernado, D. Chen *et al.*, 1995 The relationship between lineage and virulence in *Pyricularia grisea* in the Philippines. Phytopathology **85:** 443–451.

Communicating editor: P. J. Pukkila