Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 9981–9985, December 1989
Genetics

Host species-specific conservation of a family of repeated DNA sequences in the genome of a fungal plant pathogen

(rice blast/*Magnaporthe grisea*/*Pyricularia*/evolution/host–pathogen)

JOHN E. HAMER*, LEONARD FARRALL, MARC J. ORBACH, BARBARA VALENT, AND FORREST G. CHUMLEY

Central Research and Development Department, E. I. du Pont de Nemours and Company, Inc., P.O. Box 80402, Wilmington, DE 19880-0402

Communicated by Sydney Brenner, August 14, 1989 (received for review March 3, 1989)

ABSTRACT **We have identified a family of dispersed repetitive DNA sequences in the genome of *Magnaporthe grisea*, the fungus that causes rice blast disease. We have named this family of DNA sequences "MGR" for *M. grisea* repeat. Analysis of five MGR clones demonstrates that MGR sequences are highly polymorphic. The segregation of MGR sequences in genetic crosses and hybridization of MGR probes to separated, chromosome-size DNA molecules of *M. grisea* shows that this family of sequences is distributed among the *M. grisea* chromosomes. MGR sequences also hybridize to discrete poly(A)$^+$ RNAs. Southern blot analysis using a MGR probe can distinguish rice pathogens from various sources. However, MGR sequences are not highly conserved in the genomes of *M. grisea* field isolates that do not infect rice. These results suggest that host selection for a specific pathogen genotype has occurred during the breeding and cultivation of rice.**

Magnaporthe grisea is a fungal plant pathogen that shows considerable variation and diversity. For example, this species parasitizes >50 different gramineous hosts (1), yet individual isolates are restricted to one or a few grass species. Isolates of *M. grisea* that infect rice cause a devastating disease known as rice blast. This disease occurs in almost all rice-growing areas of the world, and hundreds of races have been identified in the population of rice-infecting isolates (1, 2) according to the spectrum of rice cultivars they can successfully infect. The frequent appearance of new races in the field has hindered the breeding of blast-resistant rice cultivars (2, 3).

In the laboratory, *M. grisea* isolates display variation in culture morphology (2, 4) and mating ability (4). In addition, some genetic loci appear to be highly mutable (ref. 5; B.V., unpublished results). Studies employing isozyme analysis (6) have been inconclusive in defining the nature and extent of genetic variation in this organism. Therefore, we decided to analyze the genomes of *M. grisea* rice pathogens for the presence of repetitive DNA sequences. In other organisms, repeated DNA elements have been shown to accumulate sequence polymorphisms at a high rate, presumably due to genetic drift and rearrangement (7–9). In the yeast, *Saccharomyces cerevisiae*, and in other organisms, dispersed repetitive DNA sequences participate in a wide variety of genetic rearrangements (10–12). The arrangement and structure of repetitive DNA sequences in *M. grisea* may provide clues to the mechanisms that have given rise to genetic variation in this fungal plant pathogen.

We report the identification of a family of highly polymorphic, dispersed, repeated DNA sequences ("MGR" sequences for *M. grisea* repeat) in isolates of *M. grisea* that cause rice blast disease. MGR probes do not hybridize extensively to the genomic DNA of *M. grisea* field isolates that infect grasses other than rice. However, MGR sequences are not sufficient to confer pathogenicity toward rice, because plant pathogenic strains can carry many MGR sequences without being pathogenic on rice. These results provide evidence at the DNA sequence level for pathogen genotype selection by a host species. MGR sequences appear to mark the ancestral population of *M. grisea* that led to the evolution of rice-specific pathogen genotypes.

MATERIALS AND METHODS

Strains. *M. grisea* strains are maintained in the laboratories of B.V. and F.C. (Du Pont) and J.E.H. (Purdue). Procedures for growing and crossing *M. grisea* have been described (13). *Escherichia coli* strains LE392 (14), P2392 (Stratagene), and JM105 (15) were used for the propagation of bacteriophages and plasmids.

Nucleic Acid Isolation and Blotting. DNA was isolated from purified nuclei as described by Timberlake (16). Total *M. grisea* DNA (<50 μg) was prepared by grinding 1 g of fungal mycelium in liquid nitrogen, suspending the powder in 2 ml of 50 mM Tris·HCl, pH 7.5/0.1 M EDTA/0.5% $NaDodSO_4$, and heating the suspension to 68°C. The preparation was extracted at room temperature with 2 ml of phenol/chloroform and centrifuged at 8000 × *g* for 10 min. The aqueous phase was chromatographed on a Bio-Gel P-60 (Bio-Rad) column in 10 mM Tris·HCl, pH 7.5/1 mM EDTA (TE), and nucleic acids were recovered by precipitation with isopropyl alcohol. RNA was isolated by the procedures of Timberlake (16). Poly(A)$^+$ RNA was prepared by oligo(dT)-cellulose affinity chromatography (17). DNA and RNA blots were done using Amersham Hybond-N membranes as specified by the manufacturer. Blots were probed with DNA radioactively labeled by nick-translation or by the random oligonucleotide primer method (18). A λEMBL3 (19) DNA library of strain O-137 (a rice pathogen field isolate from Hangzhou, China) was prepared and screened by plaque hybridization as described by Maniatis *et al.* (20). T4 polynucleotide kinase was used to prepare ^{32}P-labeled rRNA probes, as described (20). All nucleic acid hybridizations were conducted at high stringency, with final washes in 0.2× SSPE (1× SSPE = 0.18 M NaCl/10 mM sodium phosphate, pH 7.4/1 mM EDTA)/0.5% $NaDodSO_4$ at 65°C.

Preparation and Fractionation of Intact Chromosome-Size DNA. Protoplasts were prepared from *M. grisea* as described (21), and 1.0×10^9 protoplasts were suspended in 300 μl of STC buffer (21). The protoplast suspension was mixed with 600 μl of 1% low-melting temperature agarose in 1.0 M sorbitol/50 mM EDTA and placed on ice to harden. The agarose plugs were digested with proteinase K and washed as outlined by Orbach *et al.* (22). Contour-clamped homoge-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: MGR, *M. grisea* repeat; CHEF, contour-clamped homogeneous electric field.

*Present address: Department of Biological Sciences, Lilly Hall of Life Sciences, Purdue University, West Lafayette, IN 47906.

neous electric field (CHEF) gel electrophoresis and DNA blotting were performed essentially as described by Orbach *et al.* (22).

RESULTS

Cloning of Repetitive DNA. Recombinant clones containing repeated DNA sequences were identified in an EMBL3 library by using ^{32}P-labeled total genomic DNA as a plaque hybridization probe. We anticipated that clones containing single-copy DNA would give no signal (or a faint signal) with the genomic DNA probe, whereas clones containing repeated DNA would give an intense signal. Approximately 10% of 9000 plaques hybridized intensely to genomic DNA and not to poly(A)$^-$ RNA. Replicate filters were probed with the cloned *M. grisea* acetolactate synthase (*ILV1*$^+$) gene (F.C., unpublished results) to estimate the number of genomic equivalents that had been screened. Four clones hybridized to this single-copy probe.

We selected 40 recombinant phages that hybridized to the genomic probe and not to the poly(A)$^-$ RNA probe. We confirmed the presence of repeated DNA sequences in 32 of the λ clones. We identified 20 λ clones that carried "large" [>2 kilobases (kb) in length] repeated sequences and 12 λ clones that carried "small" (<2 kb) repeated sequences. One of the 20 λ clones with a large repeat (λMGR 12a) was labeled by nick-translation and used to probe blots of restriction digests of the 32 repeat-containing clones. Hybridization was observed to 19 of the 20 clones with large repeats and to 3 of the 12 clones with small repeats. In all cases, restriction fragments that hybridized to the λMGR 12a clone also hybridized to the genomic DNA probe. The one large repeat that did not hybridize to λMGR 12a was later shown to be an rRNA-encoding DNA (rDNA) clone that had escaped detection (J.E.H., unpublished results). This analysis yielded false positives and a clone of the rDNA genes without detecting a family of large repeated sequences other than MGR.

Structure and Distribution of MGR Sequences. We mapped restriction sites and repeated DNA regions in λMGR 12a and in four other λ clones containing MGR sequences (>2 kb). To examine the ends of MGR sequences we chose λ clones that contained at least one DNA restriction fragment that did not hybridize to the labeled genomic DNA probe. Fig. 1 shows restriction maps of the *M. grisea* genomic DNA inserts in λMGR 12a, 12b, 19, 2, and 8. The restriction fragments that hybridize to nick-translated genomic DNA are indicated. The λMGR 12a clone appears to contain a repeated sequence at least 7.5 kb in length. The remaining 5.5 kb of this λ clone do not hybridize intensely to labeled genomic DNA, and sequences in this region were confirmed by DNA hybridization analysis to be unique in the *M. grisea* genome (J.E.H. and L.F., unpublished results). Two plasmid subclones comprising the repeated sequences in λMGR 12a were constructed in pUC18. The boundaries of the *Sal* I fragments in these clones (pCB586 and pCB583) are shown below the restriction map of λMGR 12a (Fig. 1). These subclones hybridize to many different size restriction fragments in the *M. grisea* genome (see Fig. 2). Plasmids pCB583 and pCB586 were used to map the arrangement of MGR sequences in recombinant phages λMGR 12b, 19, 2, and 8, as shown in Fig. 1.

The sequence arrangement of MGR clones is strikingly heterogeneous (Fig. 1). Only λMGR 8 contains MGR sequences bounded on both sides by single-copy DNA. The repeated region in λMGR 19 appears to be interrupted by single-copy DNA. The λMGR 8 and λMGR 2 clones contain a region defined by the subclone pCB608 that is not present in λMGR 12a. The λMGR 2 and λMGR 12b clones also contain repeated DNA (not subcloned) that is not present in λMGR 12a. We conclude that MGR sequences are a highly

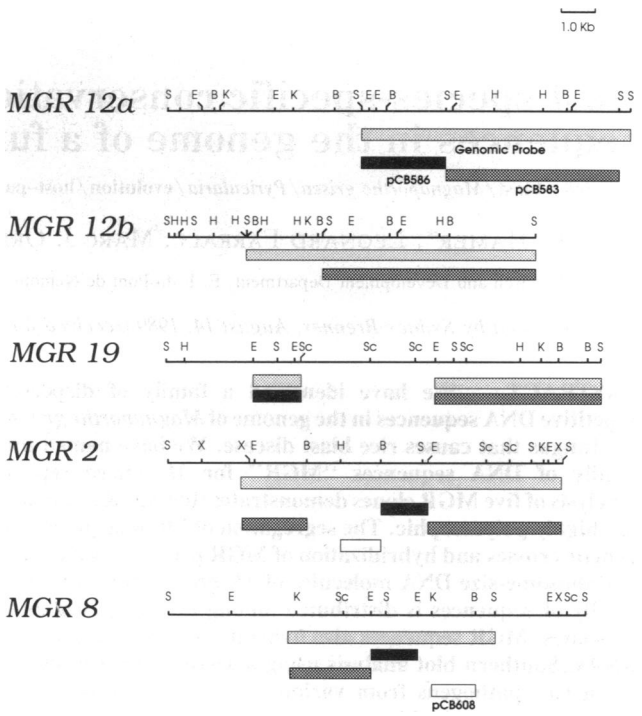

FIG. 1. Restriction maps of MGR clones. Maps of restriction sites in five MGR-containing recombinant phages are presented (thin solid horizontal lines; S, *Sal* I; E, *Eco*RI; B, *Bam*HI; K, *Kpn* I; X, *Xho* I; Sc, *Sac* I). The region that hybridizes intensely with ^{32}P-labeled genomic DNA is indicated by the light gray shaded bar. Regions of homology shared between the clones were detected by hybridization with MGR subclones: pCB586, pCB583 (black bars and dark gray bars, respectively; *Sal* I fragments of λMGR 12a), or pCB608 (open bars; *Bam*HI/*Kpn* I fragment of λMGR 8).

polymorphic repeated DNA family containing at least four distinct, but often contiguous, sequence components (represented by the pCB583, pCB586, and pCB608 subclones and the regions of λMGR 2 and 12b mentioned above).

Copy Number of MGR Sequences. The copy number of MGR sequences was determined by quantitative "dot blot" hybridization to membranes containing measured amounts of genomic DNA, using either the *ILV1*$^+$ gene or the λMGR 12a subclone, pCB583, as probes. The haploid genome size was assumed to be roughly the same as that of *Aspergillus* or *Neurospora*, ≈3.0 × 10^4 kb (23, 24). The *ILV1*$^+$ gene was determined to have a copy number of 0.8, whereas the copy number of the MGR insert in pCB583 was determined to be 37. Because the *ILV1*$^+$ gene occurs in one copy per genome (F.C., unpublished results), we normalized these results and inferred the genome size to be ≈3.8 × 10^4 kb and the MGR copy number to be about 46. This estimate for the copy number of the pCB583 MGR subclone roughly agrees with the number of bands that hybridized to pCB583 in Southern blots (Fig. 2*B*).

To examine the frequency and distribution of pCB583 and pCB586 MGR sequences in the *M. grisea* EMBL3 library, duplicate hybridization filters were prepared from each plate and probed with either pCB583 or pCB586. The number of clones containing MGR sequences was determined from autoradiographs. From five plates a total of 3912 λ plaques was screened, and a total of 424 clones (10.8%) was found to contain either one or both of the MGR regions (87 plaques hybridized to pCB586 alone; 196 plaques hybridized to pCB583 alone; 141 plaques hybridized to both probes). The frequency of recombinant clones that hybridize to pCB586 and/or pCB583 was about the same as the frequency of clones that hybridized to the labeled genomic DNA probe.

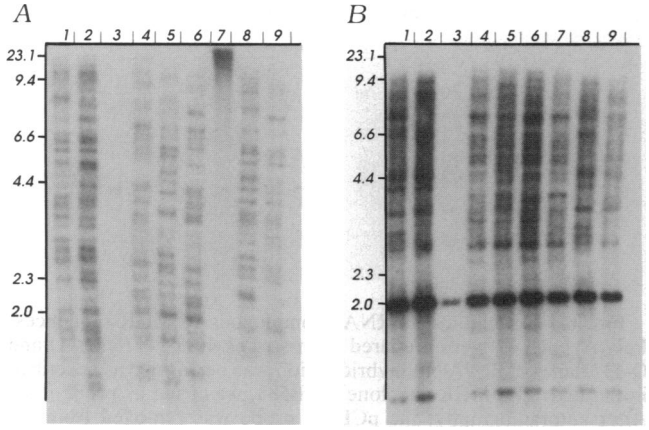

FIG. 2. Survey of MGR sequences in rice pathogenic field isolates of *M. grisea*. (*A*) DNA was digested with *Eco*RI and hybridized with pCB586. Lane 3, DNA from laboratory strain 4091-5-8, derived from Japanese nonpathogens of rice (4). The remaining lanes contained DNA from rice pathogenic field isolates as follows: 1, O-42 (Japan); 2, O-133 (Shanghai); 4, O-137 (Hangzhou); 5, O-140 (Hangzhou); 6, O-142 (Hangzhou); 7, O-155 (Beijing); 8, O-190 (Korea); 9, O-135 (Hangzhou). The DNA in lane 7 digested poorly with *Eco*RI. (*B*) Genomic DNA from the same strains as in *A* was digested with *Bam*HI, electrophoresed, blotted, and hybridized with pCB583. Lane assignments are as in *A*. (*C*) DNA from strain O-137 was digested with *Bam*HI (lane 1) or *Eco*RI (lane 2), electrophoresed, blotted, and hybridized with pCB607, which contains a 2.0-kb *Bam*HI fragment that appears to be conserved among MGR sequences. Sizes are given in kb.

Host Species-Specific Conservation of MGR Sequences in Field Isolates. The pCB586 MGR probe hybridizes to multiple *Eco*RI fragments in the genomes of rice pathogens (Fig. 2*A*). Distinctive MGR patterns are detected in each of four isolates (Fig. 2*A*, lanes 4–6 and 9) collected from different varieties of rice growing in a single field in Hangzhou, China (B.V., unpublished results). A single *Eco*RI fragment hybridizes to pCB586 in genomic DNA from strain 4091-5-8, a weeping lovegrass (*Eragrostis curvula*) pathogen that does not infect rice. The analysis of rice pathogens has been extended to include isolates from Guinea, Ivory Coast, Egypt, South Africa, Malaysia, India, Guyana, Brazil, the Philippines, and the United States (data not shown) with results similar to those shown in Fig. 2*A*. The hybridization patterns observed are sufficiently polymorphic to differentiate all of the rice pathogens we have tested.

The MGR pCB583 probe hybridizes to multiple *Bam*HI fragments in the genomes of rice pathogens (Fig. 2*B*). A highly conserved 2.0-kb *Bam*HI fragment is detected in the DNA of all strains, including strain 4091-5-8. A plasmid clone containing a copy of the 2.0-kb *Bam*HI fragment (pCB607) was obtained by isolating 2.0-kb *Bam*HI fragments from strain O-135 DNA and cloning this DNA into the plasmid vector pUC18. A Southern blot of *Eco*RI- and *Bam*HI-digested O-137 DNA was probed with pCB607 (Fig. 2*C*). As expected, the probe hybridized to the highly repeated 2.0-kb *Bam*HI fragment and to several other fragments. However, this probe detected multiple *Eco*RI fragments with no conserved components. These results suggest that MGR sequences are dispersed and consist of polymorphic and conserved DNA sequences.

We examined the distribution and conservation of MGR in field isolates of *M. grisea* that are pathogens of a variety of grasses but not pathogens of rice (Fig. 3). As with strain 4091-5-8, hybridization to pCB586 is significantly reduced in the grass pathogens as compared to rice pathogens (Fig. 3*A*). A similar result is shown for hybridization with pCB583 (Fig. 3*B*), although all strains tested show conservation of a 2.0-kb *Bam*HI fragment. The presence of numerous faintly hybridizing bands observed in longer autoradiographic exposures (J.E.H. and L.F., unpublished results) suggests the presence of a family of repeated DNA sequences distantly related to MGR in *M. grisea* nonpathogens of rice. Analysis of other *M. grisea* nonpathogens of rice isolated from a number of other geographical locations failed to detect extensive hybridization to MGR probes pCB583 and pCB586 (data not shown).

Segregation and Chromosomal Distribution of MGR Sequences. We performed a genetic cross between the weeping lovegrass pathogen 4091-5-8 and the rice pathogen field isolate O-135 (which is also a pathogen of weeping lovegrass). The segregation of MGR sequences among these progeny was determined by Southern blot analysis with the pCB586 probe (Fig. 4). The progeny analyzed in Fig. 4 are pathogens of weeping lovegrass, but none of them is a rice pathogen (B.V., unpublished results). A series of backcrosses in which progeny were mated to O-135 showed that the accumulation of MGR sequences occurred independently from the accumulation of factors that confer the ability to infect rice (B.V. and L.F., unpublished results).

The MGR bands in nine randomly chosen progeny differ in size and number (Fig. 4). Several MGR bands appear to segregate in a Mendelian manner, consistent with the location of these MGR sequences on *M. grisea* chromosomes. We did not detect any MGR bands in the progeny that were not present in the O-135 parent. In crosses between rice pathogens, we have examined random progeny without detecting anomalous segregation of MGR repeats (B.V. and L.F., unpublished results). These results suggest that MGR sequences behave as standard Mendelian markers.

We examined the hybridization of MGR sequences to chromosome-size DNA molecules separated by CHEF electrophoresis. Fig. 5*A* shows that several chromosome-size DNA molecules from *M. grisea* strains can be resolved by

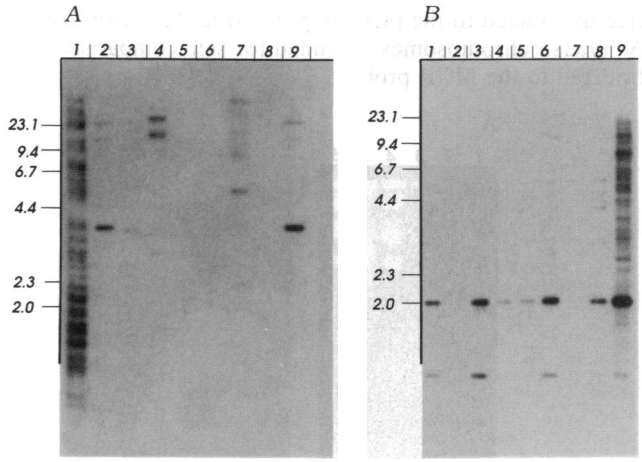

FIG. 3. Survey of MGR sequences in nonpathogens of rice. (*A*) DNA was digested with *Eco*RI and probed with pCB586. The lanes contained DNA from strains as follows: 1, O-137; 2, WGG-FA40 (finger millet, Japan); 3, 4091-5-8; 4, G-68 (napier grass, Philippines); 5, G-49 (finger millet, Japan); 6, G-48 (foxtail millet, United States); 7, G-31 (crabgrass, Japan); 8, G-26 (goosegrass, Japan); 9, G-24 (finger millet, Japan). (*B*) DNA was digested with *Bam*HI, electrophoresed, blotted, and probed with pCB583. The lanes contained DNA from the same strains as in *A*, but in the following order: 1, G-24; 2, G-26, 3, G-31; 4, G-48; 5, G-49; 6, G-68; 7, 4091-5-8; 8, WGG-FA40; 9, O-137. Equivalent amounts of DNA were loaded in all lanes, and the digested DNA was transferred completely to the hybridization membrane, as judged by ethidium bromide staining of the gel after DNA transfer. Sizes are given in kb.

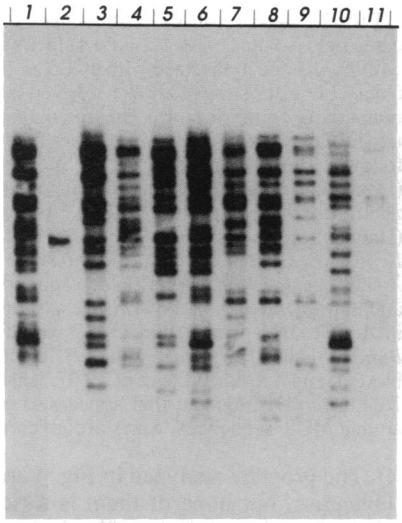

FIG. 4. Mendelian segregation of MGR sequences. Southern hybridization analysis of genomic DNA from two parental strains and nine progeny is shown. DNAs were digested with *Eco*RI and probed with pCB586. Strain O-135 (lane 1), which contains extensive MGR homology, was crossed with strain 4091-5-8 (lane 2), which contains little MGR homology. Lanes 3–11 contain DNA from random ascospore progeny of the cross: 3, 4274-R-1; 4, 4274-R-12; 5, 4274-R-21; 6, 4274-R-24; 7, 4274-R-29; 8, 4274-R-56; 9, 4274-R-62; 10, 4274-R-69; 11, 4274-R-72. None of the progeny is a pathogen of rice. Several small MGR bands in lanes 4–7 appear to be absent from the O-135 lane. However, longer exposures of the O-135 lane demonstrated the presence of these small *Eco*RI fragments in strain O-135 (results not shown).

CHEF electrophoresis. There appear to be size polymorphisms associated with the chromosomes of the strains examined. Fig. 5*B* shows that MGR sequences are present on every chromosome-size DNA molecule we have resolved in *M. grisea* rice pathogens. A single chromosome-size molecule from a *M. grisea* laboratory strain that does not infect rice hybridized to the pCB586 probe (Fig. 5*B*). None of the *N. crassa* chromosomes, included for size comparison, hybridized to the MGR probe.

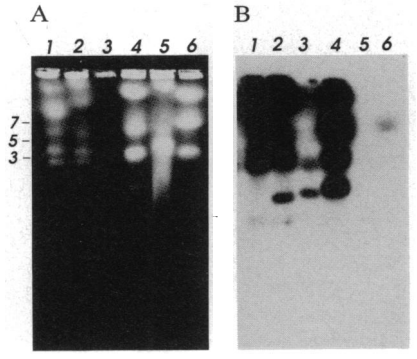

FIG. 5. Southern hybridization analysis of MGR sequences in *M. grisea* chromosomal DNAs. Chromosomal DNAs were prepared from protoplasts as described in the text. Strains examined included two rice pathogenic laboratory strains, 4375-R-6 and 4224-7-8 (lanes 1 and 2), two rice pathogenic field isolates, O-137 and O-135 (lanes 3 and 4; strain O-137 yielded poor chromosomal DNAs in this preparation), a *Neurospora crassa* strain, 74-OR23-1A (lane 5), and a *M. grisea* laboratory strain derived from nonpathogens of rice, 4136-4-3 (lane 6). (*A*) Ethidium bromide-stained CHEF gel. (*B*) Autoradiograph of the chromosomal DNAs shown in *A* after blotting and hybridizing with the MGR subclone, pCB586. The chromosomal DNAs of the *M. grisea* strains in lanes 2–4 include "minichromosomes" ≈1–2 megabases in size (difficult to see in *A*), which hybridize intensely to the pCB586 probe.

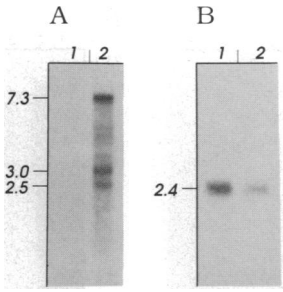

FIG. 6. Detection of mRNAs homologous to MGR sequences. Poly(A)$^+$ RNAs were prepared from strains 4091-5-8 (lanes 1) and O-137 (lanes 2). RNA blot hybridization was performed as described in the text. The MGR subclone, pCB583, was used as probe in *A*. After autoradiography, the pCB583 probe was removed from the hybridization membrane, and a plasmid carrying the *M. grisea ILV1*$^+$ gene was then used to probe the same filter, yielding the result shown in *B*.

MGR Transcription. To detect mRNAs homologous to MGR sequences, we isolated poly(A)$^+$ RNA from *M. grisea* strains O-137 and 4091-5-8 and fractionated the RNA on denaturing gels. RNA blots were probed with pCB583 (Fig. 6*A*). The same RNA blots were subsequently probed with the cloned *M. grisea ILV1*$^+$ gene (Fig. 6*B*). The MGR sequences carried by pCB583 hybridized to several poly(A)$^+$ RNAs from strain O-137: a 7.3-kb RNA and two smaller RNAs about 3.0 and 2.5 kb in length. In addition, background hybridization to O-137 poly(A)$^+$ RNAs was detected. No hybridization was observed to poly(A)$^+$ RNA isolated from strain 4091-5-8. When the probe was removed from the blot shown in Fig. 6*A*, and the blot was reprobed with the *ILV1*$^+$ gene, a single poly(A)$^+$ RNA ≈2.4 kb in length was detected. No background hybridization was observed in either lane when the *ILV1*$^+$ gene was used as a probe, demonstrating that the O-137 RNA used was not generally degraded. These results suggest that at least some MGR sequences are transcribed into mRNAs of defined lengths. The MGR sequences in the DNA of strain 4091-5-8 that hybridize to pCB583 do not appear to be transcribed in steady-state levels sufficient to be detected in these experiments. Hybridization to RNA blots using the plasmid clone containing the 2.0-kb *Bam*HI fragment described above (pCB607, Fig. 2*C*) gave results similar to those shown in Fig. 6*A* (J.E.H., unpublished results).

DISCUSSION

A family of repetitive DNA sequences exists in the genome of the rice blast pathogen, *M. grisea*. Dot blot and Southern hybridizations suggest that sequences homologous to the MGR subclone pCB583 are present in ≈46 copies per haploid genome. However, the fraction of genomic library clones that hybridize to MGR probes suggests the copy number could be somewhat higher. Although MGR sequences appear to be dispersed throughout the *M. grisea* genome and segregate independently in genetic crosses, the possibility exists that some MGR sequences may be clustered in specific genomic locations.

The MGR family is comprised of at least four different but frequently contiguous sequence components (Fig. 1; defined by subclones pCB583, pCB586, and pCB608 and by parts of λMGR 2 and 12b that have not been subcloned). Restriction mapping and Southern blot analysis demonstrate that these sequences are highly polymorphic. This sequence variation in MGR DNA fragments is presumed to have arisen by single base-pair changes and by recombination events leading to sequence rearrangements.

MGR sequences bear some resemblance to families of transposable elements found in other eukaryotic organisms. For example, the genome of *S. cerevisiae* contains ≈36 dispersed copies of a retrotransposon called Ty (25). Ty elements are ≈6 kb in length, encode 5.7-kb and 5.0-kb poly(A)$^+$ RNAs, and participate in a variety of genetic rearrangements. Like Ty elements in yeast, MGR sequences encode a large poly(A)$^+$ RNA, as well as smaller poly(A)$^+$ RNAs, suggesting that at least some MGR sequences may contain RNA polymerase II promoters and code for translatable products. The smeared pattern of RNA hybridization to the pCB583 MGR probe (Fig. 6*A*, lane 2) suggests that MGR sequences may be present in transcripts that include regions of the *M. grisea* genome adjacent to MGR insertion sites. However, we cannot rule out the possibility that MGR transcripts may be unstable during isolation or may have short half-lives. MGR sequences exhibit a much higher degree of heterogeneity and DNA sequence polymorphism than Ty elements. Nevertheless, it is possible that the MGR sequences we have studied represent deletion derivatives or partial clones of a MGR transposable element.

The MGR sequences pCB586 and pCB583 were cloned from the *M. grisea* rice pathogen, O-137. Using both of these probes we have demonstrated that MGR sequences are conserved among *M. grisea* rice pathogens isolated from different geographical areas. There is sufficient polymorphism in sequences homologous to pCB586 to demonstrate distinctive patterns of genomic *Eco*RI fragments in the DNA of individual rice pathogens. These results suggest that MGR probes may be suitable for "DNA fingerprinting" of *M. grisea* field isolates for detailed epidemiological studies. A repetitive DNA element has recently been used to characterize clinical isolates of the human fungal pathogen, *Candida albicans* (26). Similar studies in *M. grisea* may help to determine the extent of diversity and variation in field isolates from various blast prone areas.

Our results show an absolute correlation between the ability of a field isolate to infect rice and the presence of multiple bands with strong homology to MGR probes. However, MGR sequences are not likely to be determinants of pathogenicity toward rice, because genetic crosses have yielded plant pathogenic progeny that contain many MGR sequences and yet fail to infect rice (Fig. 4). We propose that the correlation between MGR sequence conservation and rice pathogenicity is due to genetic isolation and independent evolution of rice pathogens and pathogens of other grasses. *M. grisea* isolates that are nonpathogens of rice probably contain a family of repeated DNA sequences related to the MGR sequences of rice pathogens. Thus, the conservation of MGR sequences in rice pathogens from diverse locations suggests that the rice pathogens analyzed in this study are descendants from a small ancestral population. The domestication and worldwide distribution of rice appears to have provided a niche for the growth and dispersal of this ancestral rice pathogen population.

The conservation of MGR sequences closely follows the previous classification of this fungus as two species, with the name *Pyricularia oryzae* designated for rice pathogens and the name *Pyricularia grisea* designated for nonpathogens of rice (1). Our results show there has been a natural barrier to gene flow between *M. grisea* rice pathogens and nonpathogens of rice. Nevertheless, field isolates that do not infect rice can be crossed with rice pathogens to produce viable progeny that can be backcrossed to either parent (ref. 4; B.V. and L.F., unpublished results), supporting their classification as a single species.

In contrast to the high degree of polymorphism observed in MGR sequences, very little polymorphism was observed in the electrophoretic patterns of isozymes in *M. grisea* (6). Leung and Williams examined 335 field isolates of *M. grisea* from diverse geographic sites and found that 95% of these isolates were monomorphic for 16 of 18 putative enzyme loci (6). One possible explanation for this difference in polymorphism is that MGR sequences may be subjected to mutational mechanisms that produce high rates of base substitution (e.g., reverse transcription; ref. 27). Consequently, although there has been insufficient time in the evolution of *M. grisea* to accumulate neutral variation in isozyme electrophoresis patterns, sufficient time has elapsed to produce distinctive MGR hybridization patterns in each rice pathogen isolate.

Because it seems unlikely that strains capable of causing rice blast disease arose independently in various geographical areas, the large number of races among *M. grisea* rice pathogens must have emerged from the ancestral population. Furthermore, the present-day appearance of a rice-infecting variant in the general population of rice nonpathogens must be a rare or nonexistent event. The field isolate O-135 is pathogenic on rice and on weeping lovegrass, and thus could represent a pathogen of weeping lovegrass that has recently acquired the ability to infect rice. Because the DNA from strain O-135 hybridizes extensively to MGR sequences, we propose that this strain is a rice pathogen that has retained or regained an ability to infect a grass other than rice.

We thank James Sweigard, Don Ennis, and Shirley Coomber of DuPont and Morris Levy of Purdue University for their helpful suggestions and critical reading of the manuscript. We are also grateful to Richard Michelmore (University of California, Davis) for suggestions on isolating repeated DNA sequences, and we thank Sally Leong and Daniel Skinner (University of Wisconsin, Madison) for communicating published results on the identification of repeated DNA sequences in their *M. grisea* strains.

1. Ou, S. H. (1985) *Rice Diseases* (Commonwealth Mycological Institute, Surrey, UK), p. 109.
2. Latterell, F. M. (1975) in *Proceedings of the Seminar on Horizontal Resistance to the Blast Disease of Rice* (Centro Internacional de Agricultura Tropical, Cali, Colombia), Ser. CE, No. 9, pp. 199–234.
3. Ou, S. H. (1980) *Annu. Rev. Phytopathol.* **18,** 167–187.
4. Valent, B., Crawford, M. S., Weaver, C. G. & Chumley, F. G. (1986) *Iōwa State J. Res.* **60,** 569–594.
5. Hamer, J. E., Valent, B. & Chumley, F. G. (1989) *Genetics* **122,** 351–361.
6. Leung, H. & Williams, P. H. (1986) *Phytopathology* **76,** 778–783.
7. Dover, G. (1982) *Nature (London)* **299,** 111–117.
8. Ilyin, Y. V., Tchurikov, N. A., Ananiev, E. V., Rysokov, A. P., Yenikopolov, G. N., Limborska, S. A., Maleeva, N. E., Gvozdev, V. A. & Georgiev, G. P. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **42,** 959–969.
9. Strobel, E., Dunsmuir, P. & Rubin, G. M. (1979) *Drosophila* **17,** 429–439.
10. Roeder, G. S., Farabaugh, P. J., Chaleff, D. T. & Fink, G. R. (1980) *Science* **209,** 1375–1380.
11. Roeder, G. S. & Fink, G. R. (1980) *Cell* **21,** 239–249.
12. Rothstein, R., Helms, C. & Rosenberg, N. (1987) *Mol. Cell. Biol.* **7,** 1198–1207.
13. Crawford, M. S., Chumley, F. G., Weaver, C. G. & Valent, B. (1986) *Genetics* **114,** 1111–1129.
14. Murray, N. E., Brammar, W. J. & Murray, K. (1977) *Mol. Gen. Genet.* **150,** 53–61.
15. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33,** 103–119.
16. Timberlake, W. E. (1986) in *Biology and Molecular Biology of Plant-Pathogen Interactions*, NATO ASI Series H: Cell Biology, ed. Bailey, J. A. (Springer, Berlin), Vol. 1, pp. 343–357.
17. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
18. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
19. Frischauf, A. M., Lehrach, H., Poutska, A. & Murray, N. (1983) *J. Mol. Biol.* **170,** 827–842.
20. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 320–322.
21. Parsons, K. A., Chumley, F. G. & Valent, B. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4161–4165.
22. Orbach, M. J., Vollrath, D., Davis, R. W. & Yanofsky, C. (1988) *Mol. Cell. Biol.* **8,** 1469–1473.
23. Timberlake, W. E. (1978) *Science* **202,** 973–975.
24. Krumlauf, R. & Marzluf, G. A. (1980) *J. Biol. Chem.* **255,** 1138–1145.
25. Roeder, G. S. & Fink, G. R. (1983) in *Mobile Genetic Elements*, ed. Shapiro, J. A. (Academic, Orlando, FL), pp. 300–328.
26. Scherer, S. & Stevens, D. A. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 1452–1456.
27. Boeke, J. D., Garfinkel, D. J., Styles, C. A. & Fink, G. R. (1985) *Cell* **40,** 491–500.