Host Species-specific Repetitive DNA Sequence in the Genome of Magnaporthe grisea, the Rice Blast Fungus

Teruo Sone, Manabu Suto, and Fusao Tomita

Laboratory of Applied Microbiology, Department of Bioscience and Chemistry, Faculty of Agriculture, Hokkaido University, Kita-9 Nishi-9, Kita-ku, Sapporo 060, Japan

Received February 19, 1993

We cloned a repetitive sequence to show RFLPs in the genome of Magnaporthe grisea, a fungal pathogen responsible for rice blast. As the sequence was 0.8 kb in length and dispersed in the genome, it was named MGSR1 (for Magnaporthe grisea short repeat 1). MGSR1 was conserved highly in the genome of rice pathogens, but poorly in the genome of pathogens of grasses other than rice. And the RFLPs, displayed with the sequence, could distinguish between clonal lineages in rice-pathogenic isolates. The nucleotide sequence showed the presence of an internal promoter of RNA polymerase III, a 3'-poly(T), and an 8-bp direct repeat in it.

Magnaporthe grisea is a fungus that shows considerable variation and diversity in pathogenicity. Isolates of this fungus that infect rice cause a devastating disease known as a rice blast. This disease occurs in almost all rice-growing areas of the world. Hundreds of fungal races have been identified in the population of rice-infecting isolates according to the spectrum of rice cultivars that they can successfully infect. The frequent appearance of new races in the field has hindered the breeding of blast-resistant rice cultivars.

Genetic analysis of such variable pathogenicity or host specificity of this fungus has been made classically or molecular biologically by examination of enzyme polymorphism, development of fertile laboratory strains, isolation of mutants, gene cloning, transformation, gene disruption, or electrophoretic karyotyping. Recently RFLP markers of the middle repetitive DNA sequence family named MGR (for M. grisea repeat) have led to the description of interesting and useful elements specific for the rice pathogen. MGR586 in this family is commonly used for "MGF fingerprinting" because it identifies a highly polymorphic series of EcoRI restriction fragments. Romao et al. constructed a genetic map based on RFLPs of this sequence and this map is useful for further studies on genome organization and genetic analysis of this fungus. On the other hand, the genetic characterization of MGR583, another clone of MGR, found an open reading frame homologous to the human LINE family, a retropon. But these studies have been inconclusive in defining the nature and extent of genetic variation in this organism. Therefore we started to analyze the genome of this fungus using "DNA fingerprinting" with RFLP (restriction fragment length polymorphism) markers.

In this paper, we describe the finding, cloning, and genetic characterization of a new type of RFLP marker probably responsible for pathogenicity or host specificity in the genome of M. grisea.

Materials and Methods

Strains and vectors. We used six M. grisea strains, Ina 72, Guy 11, Huku 5-2, Nakayama 3-3, Iwasekadoke 2, and Yawara 4. Ina 72, Huku 5-2, and Nakayama 3-3 are rice-pathogenic field isolates from Japan. Guy 11 is a rice-pathogenic field isolate from French Guyana collected in 1979. Iwasekadoke 2 and Yawara 4 are Japanese-ginger-pathogenic and crabgrass-pathogenic isolates from Japan, respectively, which are non-rice pathogens. Escherichia coli JM109, LE392, P2392, and MV1184 were used for cloning and DNA manipulations. Plasmids pUC19, pUC118, and pUC119, and phage λEMBL3 (Stratagene, La Jolla, CA, U.S.A.) were used as vectors.

Isolation of chromosomal DNA. Each M. grisea strain was grown on prune agar (4 g/liter prune extract (Meiji Seika Kaisha, Tokyo), 1 g/liter yeast extract (Oriental Yeast Co., Ltd., Tokyo), 5 g/liter lactose, 17 g/liter agar). After the culture block (5 mm) was inoculated into 100 ml of Czapek-Dox liquid medium (Oxoid Ltd., London, UK) with 0.6% yeast extract in a 500-ml Sakaguchi flask, mycelium was grown at 27°C for 4-7 days with shaking at 130 rpm. M. grisea genomic DNA was isolated as described by Raeder et al.

Construction of genomic libraries. Two genomic libraries of strain Ina 72 were constructed as follows. Total M. grisea genomic DNA was partially digested with SmaI. The digest was ligated into the BamHI site of pUC19 or λEMBL3 with T4 DNA ligase. Recombinant clones were selected by transformation or infection of appropriate strains of E. coli.

Genomic and plaque hybridization. Equivalent amounts of genomic DNA (5 μg) was digested with restriction enzymes, electrophoresed in a 0.7% agarose gel, and transferred completely to a Hybond-N + membrane (Amersham, Buckinghamshire, UK). In Southern hybridization and plaque hybridization, ECL direct nucleic acid labeling and detection systems (Amersham) were used according to the instructions of the manufacturer.

DNA sequencing. DNA fragments were subcloned into plasmid pUC19. The DNA was sequenced by the M13 chain termination method using Sequenase ver. 2.0 (United States Biochemical Corp., Cleveland, U.S.A.) and oligonucleotide primers. Oligonucleotide primers were synthesized with a DNA synthesizer (Applied Biosystems Model 380B).

Results and Discussion

Cloning of repetitive DNA

A genomic library of strain Ina 72 using pUC19 was constructed to find RFLP (restriction fragment length polymorphism) probes that could identify host specificity of the fungi. Cloned DNA fragments on the plasmids isolated from clones, which were randomly chosen from the library, were hybridized to genomic DNA of two rice-pathogenic strains (Ina 72 and Guy 11) digested with various restriction enzymes. Some clones could detect RFLPs. One of them that was named pMG6015 hybridized
Host-specific Repetitive DNA in *M. grisea*

Conservation of repetitive DNA in isolates of *M. grisea*

Conservation of DNA fragment cloned in pMG6015 in the genome of pathogens of various grasses was examined. As shown in Fig. 1, multiple copies of pMG6015 were conserved in the rice pathogens, while very few copies were conserved in non-rice pathogens. The RFLPs between 4 rice-pathogenic isolates were observed. By hybridization patterns between 5.0 and 2.5 kbp (indicated with box in Fig. 1), these four isolates can be divided into two groups: a group consisted of Ina 72, Huku 5-2, and Nakayama 3-3 and the other of Guy11. This division was because of the location of isolation, i.e., Japan and French Guyana. Furthermore, conservation of the 2.0-kb band (indicated with arrow in Fig. 1) divided the three isolates of the Japanese-isolate group into Ina 72 (unconserved) and the others (conserved). These results suggest that this element had appeared in the genome of rice-infecting isolates of *M. grisea* after the divergence of host specificity toward rice and the other grasses, then amplified itself and dispersed in the genome of rice-pathogens individually. Thus the pMG6015 genomic insert may be suitable for "DNA fingerprinting" because it can resolve the clonal lineage relationships in rice-pathogenic isolates of *M. grisea*.

The bands in lanes 5 and 6 were observed more clearly in autoradiography with longer exposure times (data not shown). Particularly in lane 6 (crabgrass-pathogenic isolate), multiple faint bands were observed. This indicates the existence of repetitive DNA slightly homologous to the pMG6015 genomic insert in the genome of *M. grisea* crabgrass pathogens.

According to Hamer et al., a repetitive DNA in the *M. grisea* genome that was named MGR was conserved particularly in the isolates that were pathogens of rice. It is interesting that both pMG6015 and MGR are conserved specifically in rice pathogens.

Genetic characterization of pMG6015

A genomic DNA library of Ina 72 using λEMBL3 was constructed. Forty-six clones (2.63%) in a total of 1745 plaques were found to contain DNA fragments homologous to pMG6015. The DNA of 6 clones was prepared for further analysis. Restriction analysis of the clones found that the repetitive sequences dispersed at various genomic loci and had a 0.8-kb common *XhoI* fragment. It was confirmed that this fragment was present in the genomic insert of pMG6015. The Southern hybridization pattern using this fragment as a probe was the same as that

---

**Fig. 1. Conservation of DNA Fragment Cloned in pMG6015 in the Genome of Rice- or Non-rice-pathogenic Isolates of *M. grisea*.**

*M. grisea* genomic DNA was digested with BamHI and hybridized with pMG6015. The lanes 1, Ina 72 (rice pathogen); 2, Guy 11 (rice pathogen); 3, Nakayama 3-3 (rice pathogen); 4, Huku 5-2 (rice pathogen); 5, Iwasekadoge 2 (Japanese ginger pathogen); 6, Yawara 4 (crabgrass pathogen). Clonal lineage-specific RFLP region (box) and RFLP band (arrow) are indicated.

---

**Fig. 2. Restriction Map of pMG 6015-3.**

Basic map of all genomic insert of the clone (upper) and detailed map of repetitive DNA-including region (lower). λEMBL3 (open bars), single-copy DNA (dark-gray shaded bars), and pMGT6-1 genomic insert (striped bar) are indicated.

---

Conservation of repetitive DNA in isolates of *M. grisea*

Conservation of DNA fragment cloned in pMG6015 in the genome of pathogens of various grasses was examined. As shown in Fig. 1, multiple copies of pMG6015 were conserved in the rice pathogens, while very few copies were conserved in non-rice pathogens. The RFLPs between 4 rice-pathogenic isolates were observed. By hybridization patterns between 5.0 and 2.5 kbp (indicated with box in Fig. 1), these four isolates can be divided into two groups; a group consisted of Ina 72, Huku 5-2, and Nakayama 3-3 and the other of Guy11. This division was because of the location of isolation, i.e., Japan and French Guyana. Furthermore, conservation of the 2.0-kb band (indicated with arrow in Fig. 1) divided the three isolates of the Japanese-isolate group into Ina 72 (unconserved) and the others (conserved). These results suggest that this element had appeared in the genome of rice-infecting isolates of *M. grisea* after the divergence of host specificity toward rice and the other grasses, then amplified itself and dispersed in the genome of rice-pathogens individually. Thus the pMG6015 genomic insert may be suitable for "DNA fingerprinting" because it can resolve the clonal lineage relationships in rice-pathogenic isolates of *M. grisea*.

The bands in lanes 5 and 6 were observed more clearly in autoradiography with longer exposure times (data not shown). Particularly in lane 6 (crabgrass-pathogenic isolate), multiple faint bands were observed. This indicates the existence of repetitive DNA slightly homologous to the pMG6015 genomic insert in the genome of *M. grisea* crabgrass pathogens.

According to Hamer et al., a repetitive DNA in the *M. grisea* genome that was named MGR was conserved particularly in the isolates that were pathogens of rice. It is interesting that both pMG6015 and MGR are conserved specifically in rice pathogens.

Genetic characterization of pMG6015

A genomic DNA library of Ina 72 using λEMBL3 was constructed. Forty-six clones (2.63%) in a total of 1745 plaques were found to contain DNA fragments homologous to pMG6015. The DNA of 6 clones was prepared for further analysis. Restriction analysis of the clones found that the repetitive sequences dispersed at various genomic loci and had a 0.8-kb common *XhoI* fragment. It was confirmed that this fragment was present in the genomic insert of pMG6015. The Southern hybridization pattern using this fragment as a probe was the same as that
was selected for sequencing of the nucleotides of MGSRI, the new probe for "DNA fingerprinting."

The nucleotide sequence of MGSRI1 in pMG76-1 is shown in Fig. 3. It had two internal promoter regions of RNA polymerase III, a poly(T) tract at its 3′end, and a direct repeat of 8 nucleotides. These characteristics indicate that MGSRI1 would be a retroposon with an RNA polymerase III promoter.

The retroposons with RNA polymerase III promoter are ordinarily considered to be SINEs (short interspersed elements), which is a transposable element family in higher eucaryotes. SINEs are short retroposons usually having oligo(A) or oligo(T) in their 3′ends, 7 to 21 bp direct repeats, and internal RNA polymerase III promoters. Several mammalian SINEs closely resemble specific tRNAs. 5 Our MGSRI also has a poly(T) and a direct repeat, but they are different from SINEs. In every SINE, the direct repeat lies downstream of 3′poly(T), but in MGSRI, it lies upstream of poly(T). Upon GenBank (R71.0 March, 1992), nucleotide sequence database, we could not find any tRNAs nor tRNA-based SINEs with significant homology with MGSRI1. Experiments to give evidence of the transposition of MGSRI1 are now in progress.

RFLP analysis has been very useful to resolve the complexity of the genome of many organisms. In studies on the pathogenicity of M. grisea, "DNA fingerprinting" using MGR probes called "MGR fingerprinting" displayed its potential for resolving the genome dynamics in M. grisea. 5 Our MGSRI1 also will contribute to examine the variability in host specificity and pathogenicity of the fungi as a probe for "DNA fingerprinting" of M. grisea rice-pathogenic isolates.

Acknowledgments. We thank Dr. L. Notteghem for providing Guy 11, Dr. S. Kiyosawa for providing us several Japanese isolates of M. grisea. We also thank Biopolymer Analysis Laboratory, Faculty of Agriculture, Hokkaido University for the advice of manipulation of DNA synthesizer.

References