Fungal Genetics and Biology 45 (2008) 1340–1347

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Development of microsatellite markers and construction of genetic map in rice blast pathogen *Magnaporthe grisea*

Yan Zheng [a], Geng Zhang [a], Fucheng Lin [b], Zonghua Wang [c], Gulei Jin [a], Long Yang [a], Ying Wang [a], Xi Chen [a], Zhaohua Xu [a], Xiangqian Zhao [a], Hongkai Wang [a], Jianping Lu [a], Guodong Lu [b], Weiren Wu [a,d,*]

[a] *Department of Agronomy, College of Agriculture & Biotechnology, Zhejiang University, Hangzhou, Zhejiang 310029, China*
[b] *Institute of Biotechnology, College of Agriculture & Biotechnology, Zhejiang University, Hangzhou, Zhejiang 310029, China*
[c] *College of Plant Protection, Fujian Agriculture & Forestry University, Fuzhou, Fujian 350002, China*
[d] *College of Life Sciences, Fujian Agriculture & Forestry University, Fuzhou, Fujian 350002, China*

ARTICLE INFO

Article history:
Received 23 October 2007
Accepted 16 July 2008
Available online 24 July 2008

Keywords:
Magnaporthe grisea
Simple sequence repeat (SSR)
Molecular marker
Genetic map

ABSTRACT

Magnaporthe grisea is the most destructive fungal pathogen of rice and a model organism for studying plant–pathogen interaction. Molecular markers and genetic maps are useful tools for genetic studies. In this study, based on the released genome sequence data of *M. grisea*, we investigated 446 simple sequence repeat (SSR) loci and developed 313 SSR markers, which showed polymorphisms among nine isolates from rice (including a laboratory strain 2539). The number of alleles of each marker ranged 2–9 with an average of 3.3. The polymorphic information content (PIC) of each marker ranged 0.20–0.89 with an average of 0.53. Using a population derived from a cross between isolates Guy11 and 2539, we constructed a genetic map of *M. grisea* consisting of 176 SSR markers. The map covers a total length of 1247 cM, equivalent to a physical length of about 35.0 Mb or 93% of the genome, with an average distance of 7.1 cM between adjacent markers. A web-based database of the SSR markers and the genetic map was established (http://ibi.zju.edu.cn/pgl/MGM/index.html).

1. Introduction

Rice blast is the most destructive disease of rice caused by the fungal pathogen *Magnaporthe grisea* (Hebert) Barr [(*Pyricularia grisea*, Sacc) *Pyricularia oryzae*, Cavara]. *M. grisea* can infect various parts of rice including leaf, stem, spike and root (Sesma and Osbourn, 2004). *M. grisea* is highly variable and can parasitize a wide spectrum of hosts (>50 grass species; Sweigard et al., 1995). *M. grisea* has been regarded as a 'principal model organism' for the study of molecular mechanisms of fungal diseases and host–pathogen interactions in plants (Valent, 1990; Dean, 1997; Martin et al., 2002; Dean et al., 2005).

Molecular markers and genetic maps are useful tools for genetic studies such as gene mapping, genome comparison and genetic diversity analysis. Nagakubo et al. (1983) first reported a genetic map of *M. grisea* consisting of 11 auxotrophic and fungicide resistance markers. The use of molecular markers to construct genetic maps in *M. grisea* began in late 1980s. Early maps of *M. grisea* were mainly based on restriction fragment length polymorphism (RFLP) markers (Skinner et al., 1990, 1993; Romao and Hamer, 1992; Budde et al., 1993; Sweigard et al., 1993; Farman and Leong, 1995; Nitta et al., 1997). Because the technology of RFLP analysis is complicated and laborious, application of these maps was limited.

Simple sequence repeat (SSR) or microsatellite is a PCR-based molecular marker, which has many advantages including generally codominant, highly reproducible, highly polymorphic and abundant in animal, plant and microbe genomes (Maroof et al., 1994). SSR has become the most popular molecular marker system and has been extensively utilized in genetic mapping and gene tagging (Wang et al., 2005), genetic diversity studies (Struss and Plieske, 1998; Tenzer et al., 1999) and pedigree analysis (Rongwen et al., 1995). Karaoglu et al. (2004) suggested that there are at least 11,000 SSRs with lengths of 10 bp or more in *M. grisea*. Hence, SSRs have great potential to be exploited as molecular markers in *M. grisea*. In recent years, efforts have been made to develop SSR markers for constructing genetic maps in *M. grisea*. (Brondani et al., 2000) first developed 24 SSR markers for genetic analysis. Kaye et al. (2003) integrated 23 SSR markers into the genetic map constructed by Nitta et al. (1997). More recently, Wang et al. (2005) constructed a genetic map consisting of 121 SSR, 9 SCAR and 1 RAPD markers for mapping avirulence genes. Ma et al. (2006) also developed 121 SSR markers and mapped an avirulence gene. However, both Wang et al. (2005) and Ma et al. (2006) did not published the primer sequences of their SSR markers and the number of

* Corresponding author. Address: Department of Agronomy, College of Agriculture & Biotechnology, Zhejiang University, Hangzhou, Zhejiang 310029, China. Fax: +86 571 86971910.
E-mail address: wuwr@zju.edu.cn (W. Wu).

1087-1845/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.fgb.2008.07.012

available SSR markers in *M. grisea* to date still cannot meet the need of genetic studies.

Recently, a draft of *M. grisea* genome sequence has been released (Dean et al., 2005). This provides us with an opportunity to exploit a large number of new SSR markers in *M. grisea*. In this study, we developed a set of SSR markers using the available *M. grisea* genome sequence data and constructed a genetic map of *M. grisea* consisting of 176 SSR markers. We also established a web-based database of the SSR markers and genetic map, which will facilitate gene mapping and map-based gene cloning in *M. grisea*.

2. Materials and methods

2.1. Identification of SSRs and design of primers

Genome sequence data of *M. grisea* (versions 2 and 5) were downloaded from website http://www.broad.mit.edu/annotation/genome/magnaporthe_grisea. Searching of SSRs was conducted mainly based on the data of version 2 before the release of version 5. Intact SSRs with at least five repeats for dinucleotide and trinucleotide types or four repeats for tetranucleotide type were identified using program SSRIT (http://www.gramene.org/gramene/searches/ssrtool). SSRs of mononucleotide repeat type with at least 20 nucleotides in length were identified by manual work. A set of SSRs approximately evenly distributed in the physical map were selected. In order to increase the possibility of obtaining polymorphic markers, we usually chose the SSR with the largest number of repeats in spite of its repeat motif when there were several SSRs available within a target region. A pair of primers bracketing each SSR was designed using program Primer3 (http://www.frodo.wi.mit.edu/cgi-bin/primer3/). The designed primers were evaluated by electronic PCR (e-PCR) (Schuler, 1997) on the whole genome sequence (versions 2 and 5). Only those primer pairs that produced a unique sequence as expected in the e-PCR were selected as candidate SSR markers for experiments. For regions where adjacent SSR markers were distantly spaced (>30 cM) in the genetic map (see Section 2.5), additional SSR markers within those regions were developed. All primers used in the experiments were synthesized by Shanghai Sangon Biological Engineering & Technology Company.

2.2. Detection of polymorphisms of candidate SSR markers

Nine *M. grisea* isolates from rice (including a laboratory isolate) were used for detecting the polymorphisms of the candidate SSR markers (Table 1). The polymorphism level of each SSR marker was indicated by its polymorphism information content (PIC) value calculated with the following formula (Anderson et al., 1993):

$$\mathrm{PIC} = 1 - \sum_{i=1}^{n} P_i^2$$

where p_i is the frequency of the ith allele and n is the number of alleles.

Table 1
M. grisea strains isolated from rice used for analyzing SSR polymorphisms

Isolate	Source
Guy11	French Guyana, France
2539	Laboratory strain (Leung et al., 1988)
81278	Fujian, China
97045B	Fujian, China
8-7-1	Sichuang, China
8-65-1	Sichuang, China
98-277-1	Zhejiang, China
99-06-2	Zhejiang, China
96-4-1a	Yunnan, China

2.3. Construction of mapping population

Hybridization between isolates Guy11 and 2539 was performed following the method of Valent et al. (1991). Hundreds of asci produced by the hybrid were mixed in 1 ml sterilized water and placed at 25 °C for 2–3 h to allow ascospores released. The released ascospores were spreaded on 4% water agar medium and then individual ascospores were randomly isolated under a 100× microscope using a fine glass needle. Since there were a large number of asci mixed together, it would be very unlikely to isolate sister ascospores from the same ascus. This assumption was verified posteriorly, since we did not find two progeny with the same multilocus genotype. All the ascospores isolated were cultured separately and allowed to develop into strains. A total of 299 strains were obtained, among which 228 and 71 were gray and buff, respectively. It has been known that buff strains have lost the ability of infecting rice (Chumley et al., 1990). So, we did not use these buff strains for genetic mapping. By inoculating the 228 gray strains to a rice cultivar CO39, we found that 121 strains could infect CO39 and 107 strains could not, suggesting that there might be a pair of virulent and avirulent alleles segregating in the population. This was consistent with previous reports that an avirulence gene *AVR1-CO39* was mapped on chromosome 1 (Smith and Leong, 1994) and further cloned by positional cloning (Farman and Leong, 1998) based on the same cross, with the virulent and avirulent alleles coming from Guy11 and 2539, respectively. PCR analysis confirmed that the *AVR1-CO39* genotypes of the 228 strains were consistent with the inoculation phenotypes. According to the genetic structure of the population, we decided to use all the 228 gray strains for constructing the map of chromosome 1, but only selected the 121 virulent strains that did not carry the *AVR1-CO39* allele from the parental isolate 2539 for the map construction of other chromosomes. The selection would not influence the linkage analysis of chromosomes 2–7 because they all are independent of the *AVR1-CO39* locus. This was verified later as most of the markers on these chromosomes segregated normally among the virulent strains (see Section 3.2).

2.4. Genomic DNA extraction and PCR

Genomic DNA of each strain was extracted following the method of Talbot et al. (1993). PCR for SSR analysis was performed in a 15 μl reaction mixture containing 20–50 ng genomic DNA, 1× PCR Buffer (10 mM Tris–HCl, pH 8.4, 50 mM KCl, 1.5 mM $MgCl_2$), 200 μM of each dNTP, 0.5 μM each of forward and reverse primers, and 1 U of *Taq* polymerase. All SSR primers were tested with a routine PCR program: 5 min at 94 °C; 30 cycles of 30 s at 94 °C, 30 s at 55 °C, and 1 min at 72 °C; and 5 min at 72 °C for final extension. We used 6% non-denaturing PAGE (300 V, 2 h) for separating PCR products and silver staining for visualizing DNA bands following the method of Xu et al. (2002) with minor modifications.

PCR for detecting gene *AVR1-CO39* was performed following Tosa et al. (2005) with a pair of specific primers AVF1-2 (5′-TGCCG CATTTTGCTAACC-3′) and AVR1-2 (5′-GCGAATCCATAGACAAGGA C-3′). PCR for determining mating type was performed following Wang et al. (2004) using two pairs of specific primers: MAT1-1F (5′-TCAGCTCGCCCAAATCAACAAT-3′) and MAT1-1R (5′-ACTCAAG ACCCGGCACGAACAT-3′) for mating type *MAT1-1*, and MAT1-2F (5′-GAGTTGCCTGCCCGCTTCTG-3′) and MAT1-2R (5′-GGCTTGGTC GTTGGGGATTGT-3′) for mating type *MAT1-2*. The PCR products were electrophoresed in 1% agarose gel and stained with ethidium bromide.

2.5. Linkage analysis

Linkage analysis was performed using program MAPMAKER/EXP 3.0 (Lander et al., 1987). Kosambi's mapping function was cho-

sen. Three anchored markers for each chromosome were selected by referring to the physical map (version 5). All other markers were assigned to the chromosomes using ASSIGN command with a LOD threshold of 3.0. For each chromosome, those markers with known positions in the physical map (version 5) were selected to make a framework with the same marker order as that in the physical map; and then TRY command was used to determine the locations of remaining markers in the framework. After the marker order was determined, the distances between adjacent markers were calculated using MAP command with ERROR DETECTION ON. The genetic map was drawn using program Mapdraw v2.1 (Liu and Meng, 2003).

3. Results

3.1. Polymorphisms of SSR markers

A total of 446 primer pairs were designed and tested (Supplementary Table 1), of which 418 pairs (MGM1–MGM418) were designed based on the genome sequence data of version 2 and 28 pairs (MGM419–MGM454) were designed based on version 5. Four hundred and twenty-five primer pairs amplified PCR products with expected sizes in at least 3 of the 9 isolates analyzed (345 pairs amplified products in all the 9 isolates, 49 pairs did in 8 isolates, 16 did in 7, 9 in 6, 4 in 5, 1 in 4 and 2 in 3, respectively), of which 313 (73.6%) showed polymorphisms and therefore could be used as SSR markers. The percentage of polymorphic SSRs observed here is very close to that (73%) reported by Kaye et al. (2003). The number of alleles detected per locus ranges 2–9 with an average of 3.3, which is almost the same as the estimate (3.4) reported by Adreit et al. (2007). The PIC value of each marker varies 0.20–0.89 with an average of 0.53. According to the criteria proposed by Botstein et al. (1980), 168 (53.7%) of the SSR markers are highly informative (PIC > 0.5); 97 (31.0%) are reasonably informative (0.25 < PIC < 0.5); only ~15% are slightly informative (PIC < 0.25).

Since we chose SSR loci mainly according to their repeat numbers instead of their repeat motifs, various types of SSRs (except mononucleotide type) were nearly randomly selected. Therefore, the numbers of various SSR types could roughly reflect their relative abundances. The result shows that tetranucleotide type is much fewer than dinucleotide and trinucleotide types in the *M. grisea* genome (Table 2). This is consistent with the result of Karaoglu et al. (2004).

We used three indexes (percentage of polymorphic SSRs, number of alleles per locus and PIC value) to indicate SSR polymorphism level (SPL). Statistical analyses on these indexes showed that SPL depends on SSR motif and SSR sequence length. There is a general trend that SPL decreases as the length of repeat motif increases; tetranucleotide type has the lowest polymorphism level (Table 2). This is consistent with the result reported by Wang et al. (2005). SPL also shows a negative relationship with the G/C content of SSR sequence in general; the motif GGC/CCG is the least polymorphic in the trinucleotide type (Table 2). In contrast, SPL is positively correlated with the number of repeats or the length of SSR sequence (Table 3). No significant difference in SPL was detected among different chromosomes (data not shown), suggesting that SPL is randomly distributed in the genome.

It is necessary to mention that among the SSR markers we developed, there is a marker (named MGM269) based on a sequence of 10 repeats of a nine-nucleotide motif CTTTTTTTTT (Supplementary Table 1). The marker showed a high polymorphism level (number of alleles = 5; PIC = 0.741) among the nine isolates tested. Although the marker does not follow the general definition of SSR, we still regard it as a SSR marker.

3.2. Genetic map

One hundred and eighty-one SSR markers that showed codominant polymorphisms between Guy11 and 2539 and could be unambiguously detected by non-denaturing PAGE were selected for the construction of genetic map. The map acquired consists of 176 SSR markers distributed on seven chromosomes, covering a total length of 1247 cM, equivalent to a physical length of about 35.0 Mb or 93% of the genome, with an average distance of 7.1 cM between adjacent markers (Fig. 1). The number of markers on each chromosome, length of each chromosome and average distance between adjacent markers on each chromosome varies 17–37, 124.7–256.7 cM and 5.67–7.85 cM, respectively (Fig. 1). The variation of the average distance between adjacent markers among the chromosomes is not statistically significant.

Most of the SSR markers with known positions in the physical map (version 5) could be assigned to the expected chromosomes and reasonably mapped in the given order. However, five markers (MGM86, MGM358, MGM121, MGM258 and MGM259) could not be located in the genetic map. In addition, four markers from chromosome 2 (MGM200 on contig5.734; MGM201, MGM445 and

Table 2
Polymorphisms of various SSRs in *M. grisea*

Repeat type/motif[a]	SSR no.	Polymorphic SSR[b]			
		No.	%	Allele no.	PIC
Mono-	47	47	100.0 Aa	3.17 Bb	0.538 Aa
Di-	147	127	86.4 Bb	3.69 Aa	0.568 Aa
Tri-	188	113	60.1 Cc	3.11 Bc	0.508 Aa
Tetra-	43	26	60.5 Cc	2.92 Cd	0.422 Bb
AT/TA (0)	29	26	89.7 Aa	4.15 Aa	0.635 Aa
AG/CT (1)	61	53	86.9 Aa	3.74 Aab	0.565 Aab
AC/GT (1)	57	48	84.2 Aa	3.38 Ab	0.534 Ab
AAT/ATT (0)	5	4	80.0 Aa	3.50 ABab	0.620 Aa
AAC/GTT (1)	22	16	72.7 Aa	4.00 ABab	0.631 Aa
AAG/CTT (1)	10	7	70.0 ABa	4.14 Aa	0.602 Aa
ATG/CAT (1)	81	51	63.0 Aa	2.78 ABb	0.468 ABa
ACC/GGT (2)	10	8	80.0 Aa	2.88 ABb	0.472 ABab
AGC/GCT (2)	47	25	53.2 ABab	3.00 ABab	0.495 ABab
GGC/GCC (3)	10	2	20.0 Bb	2.00 Bb	0.321 Bb

[a] The number in brackets indicates the number of G/C in the motif.
[b] Capital letters (A, B, C) and small letters (a, b, c, d) indicate significance of difference at 1% and 5% levels, respectively, according to *t*-test (for percentage of polymorphic SSRs) or Duncan's new multiple range test (for number of alleles and PIC value).

Table 3
Correlation (*r*) between polymorphism level (allele number or PIC value) and repeat number in various SSR types in *M. grisea*

Repeat type	Polymorphic SSR no.	Range of repeat no.	Allele no.		PIC	
			r	*P*-value	*r*	*P*-value
Mono-	47	22–59	0.18	0.214	0.37	0.011
Di-	127	5–68	0.49	4.70×10^{-9}	0.46	4.53×10^{-8}
Tri-	113	5–37	0.43	2.40×10^{-6}	0.33	3.85×10^{-4}
Tetra-	26	4–48	0.91	8.08×10^{-11}	0.72	2.26×10^{-5}

MGM452 on contig5.736, which is adjacent to contig5.734) and one from chromosome 7 (MGM88 on contig5.173) were mapped to chromosome 4 (Fig. 1). Of these five markers, three (MGM88, MGM200 and MGM201) were mapped to one chromosome end; the other two (MGM445 and MGM452) were mapped to intra-chromosomal region. It is noticeable that although MGM201, MGM445 and MGM452 are located on the same contig in the physical map, MGM201 is far separated from MGM445 and MGM452 but close to MGM200 in the genetic map.

There were 11 markers with unknown positions in the physical map (version 5) being located in our genetic map. They were mapped to six chromosomes, with one each to chromosomes 1, 2, 3 and 6, two to chromosome 5 and five to chromosome 4 (Fig. 1). Four of the markers were mapped to chromosome ends and seven to intra-chromosomal regions. Interestingly, six of these markers (including MGM423 on chromosome 1, MGM431 on chromosome 2, MGM430, MGM426 and MGM427 on chromosome 4, and MGM263 on chromosome 6) happened to be located within intervals where gaps existed in the physical map (Fig. 1), suggesting that their corresponding contigs might fall in those gaps.

The avirulence gene *AVR1-CO39* was mapped between markers MGM23 and MGM311 on chromosome 1 (Fig. 1). This is consistent

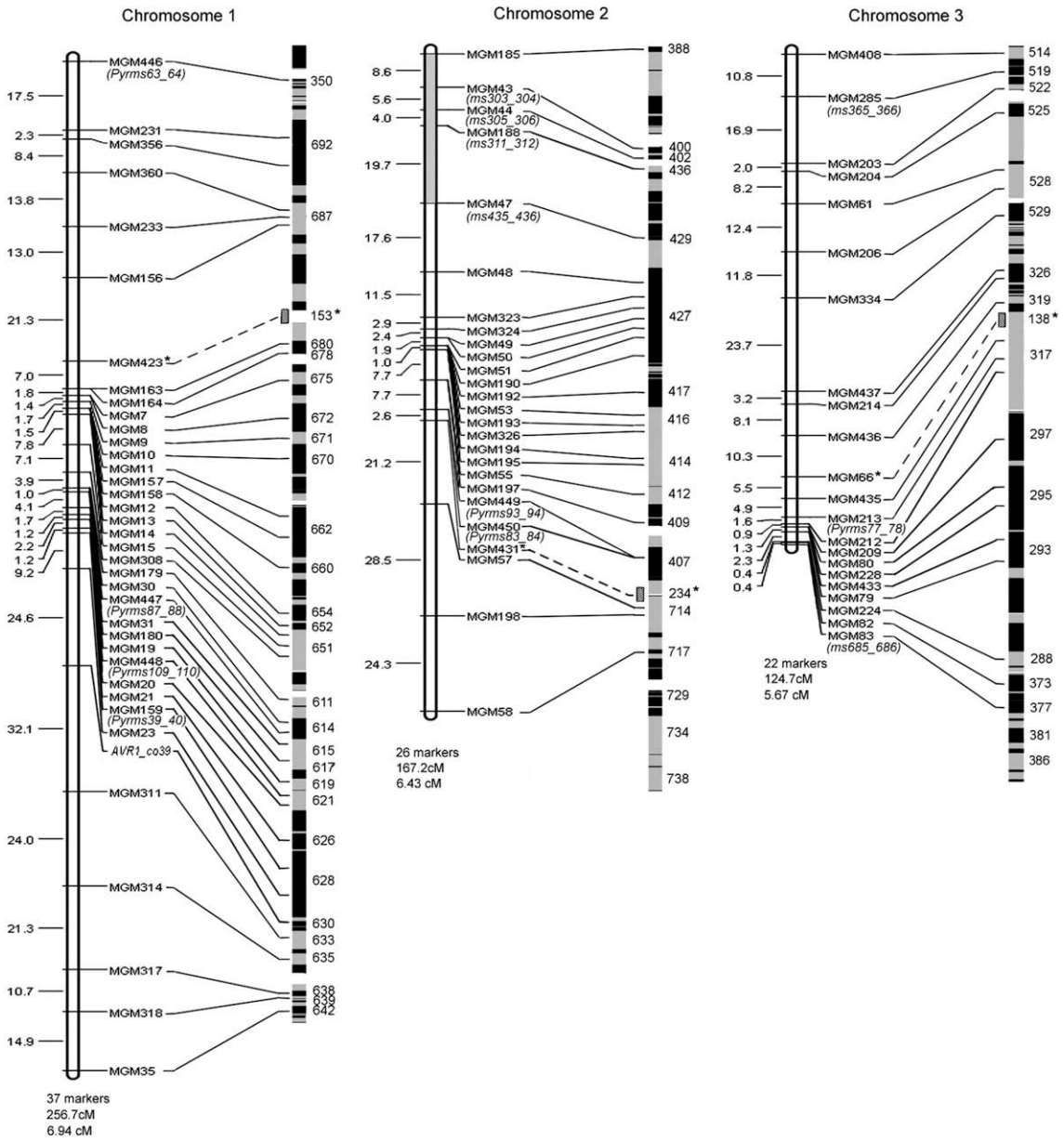

Fig. 1. Genetic map of *M. grisea* constructed based on the order of physical map (version 5) and its comparison with the physical map. For each chromosome, the genetic map is on the left and the physical map is on the right. In the genetic map of each chromosome, names of SSR markers are shown on the right side and distances (cM) between adjacent markers are shown on the left side; the number of markers, chromosome length (cM) and average distance between adjacent markers (cM) are shown under the map. Markers typed in italic are from Kaye et al. (2003) (prefix 'Pyrms') or Wang et al. (2005) (prefix 'ms'), equivalent to the 'MGM' markers directly above them. In the physical map, the black or light gray boxes represent contigs, of which the names are shown on the right side (note: the full name of a contig is in the form of 'contig5.number'; the prefix 'contig5.' is omitted in the figure). The disconnections in the physical map represent gaps. Markers with unknown physical positions or mapped to altered positions (not expected according to the physical map) are marked with asterisk or octothorpe; their corresponding contigs are represented by dark gray boxes in the physical map and the names of the contigs are also marked with asterisk or octothorpe accordingly. Markers and corresponding contigs are all connected with solid or dashed lines. The contig '2.1177' corresponding to marker MGM171 in chromosome 5 is from the genome sequence data of version 2. Hot spots of segregation distortion are indicated by either light gray (on chromosome 2, biased to 2539) or dark gray (on chromosomes 6, biased to Guy11) in the genetic map.

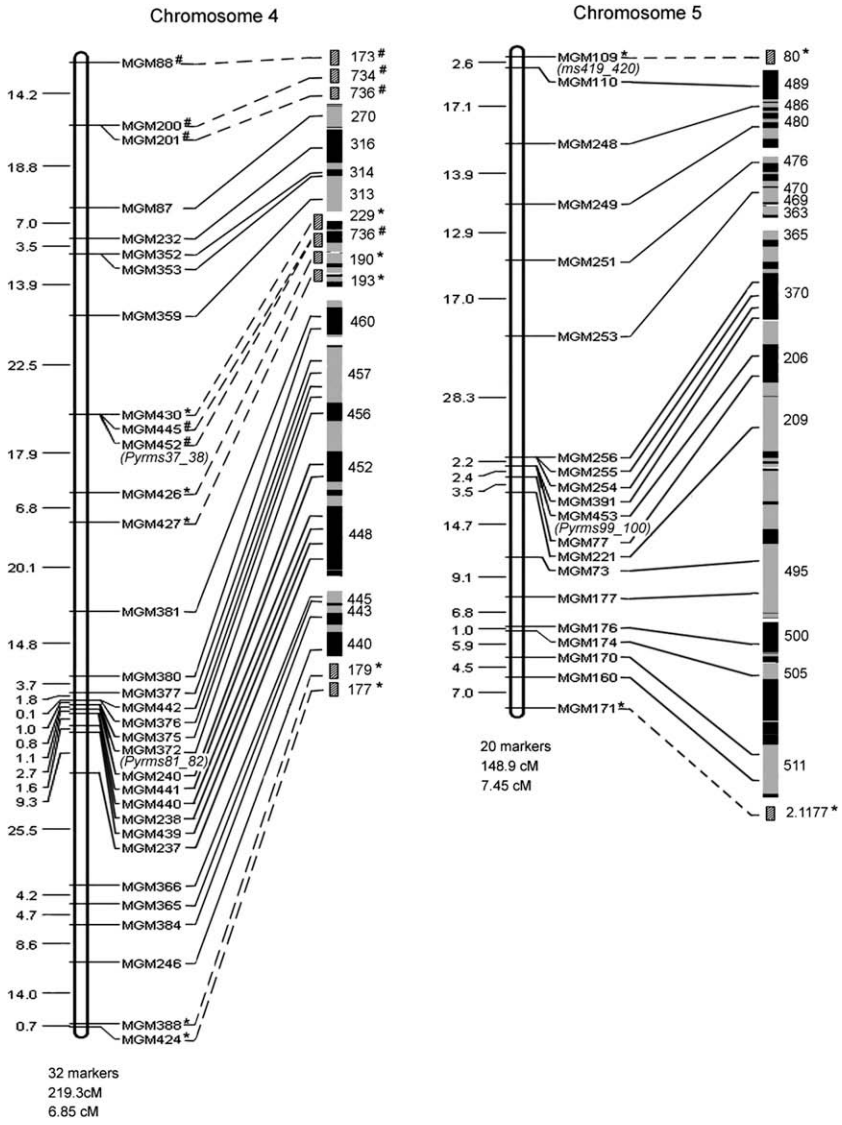

Fig. 1 (*continued*)

with the results obtained by Smith and Leong (1994) and Farman and Leong (1998)). The mating type gene *MAT* was mapped between markers MGM297 and MGM298 on chromosome 7 (Fig. 1). This is consistent with the result reported by Nitta et al. (1997).

An overall comparison between the genetic map and the physical map (Fig. 1) shows that 1 cM is approximately equivalent to 28 kb on average in the *M. grisea* genome in this experiment. The average ratio of physical distance to genetic distance (28 kb/cM) might be somewhat underestimated because the lengths of gaps in the physical map are not considered. Nevertheless, the result still suggests that the genetic recombination frequency is generally quite high in *M. grisea*. However, the recombination frequency is not evenly distributed but dramatically fluctuates along the genome. Many recombination hot spots can be identified in the map, such as regions MGM360–MGM156 and MGM23–MGM314 (containing the *AVR1-CO39* locus) on chromosome 1, MGM57–MGM198 on chromosome 2, MGM265–MGM266 on chromosome 6, and MGM417–MGM418 on chromosome 7, where 1 cM is approximately equivalent to 9.4, 4.9, 7.6, 3.3 and 2.4 kb, respectively. There are also a number of recombination cold spots (Table 4). The 'coldest' spot would be the region between MGM212 and MGM83 on chromosome 3, where 1 cM is approximately equivalent to 515 kb.

Most markers showed normal segregation ratio (1:1) as expected in the population. However, 26 (14.4%) markers (including three unmapped markers MGM121, MGM258 and MGM259 of chromosome 5; see above) exhibited distorted segregation ratio at the 0.01 significance level, with 16 markers biased to parent Guy11 and 10 markers biased to parent 2539. These distorted markers are mainly distributed on chromosomes 2 and 6. The distorted makers on chromosome 2 are all biased to 2539, while those on chromosome 6 are all biased to Guy11.

3.3. Database

We have established a web-based database named MGM, which is accessible at website http://ibi.zju.edu.cn/pgl/MGM/index.html. A list of all of the SSRs and detailed information of each SSR, including position, primers, repeat motif, expected PCR product length, anneal temperature, number of alleles detected in the nine *M. grisea* isolates and PIC values, are provided. In addition, the genetic map (Fig. 1) was displayed in a setting

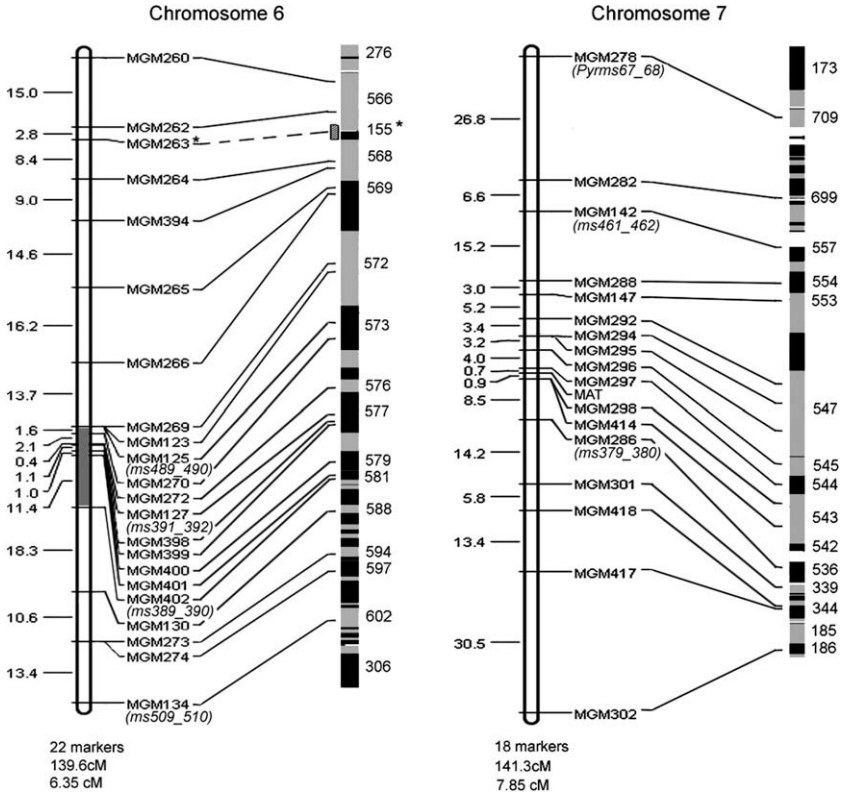

Fig. 1 (*continued*)

Table 4
Recombination cold spots on each chromosome

Chromosome	Average PD/GD[a] (kb/cM)	Recombination cold spot	GD (cM)	PD (Mb)	PD/GD[b] (kb/cM)
1	29.0	MGM163–MGM14	6.4	2.52	>393.8
		MGM179–MGM159	11.4	1.65	>144.7
2	29.2	MGM323–MGM195	8.2	1.58	192.7
3	43.9	MGM212–MGM83	5.3	2.73	>515.1
4	20.0	MGM377–MGM439	9.1	1.34	147.3
5	35.5	MGM256–MGM221	8.1	1.11	137.0
6	28.6	MGM269–MGM401	6.2	1.65	266.1
7	25.3	MGM292–MGM414	12.2	0.99	81.1

[a] PD, physical distance; GD, genetic distance.
[b] The symbol '>' indicates that the values could be higher because there are gaps within the regions in the physical map.

allowing the identification of the detailed information on each specific marker.

4. Discussion

Most (73.6%) SSRs we have identified in *M. grisea* genome are polymorphic among a set of nine *M. grisea* isolates, 84.7% of which are highly to reasonably informative. This means that more than 60% of *M. grisea* SSRs could be exploited as informative genetic markers. This proportion is likely overestimated as we have selected for mapping SSRs with a high number of repeats. However, we can still conclude that the polymorphism level of SSRs in *M. grisea* is high enough for being used as an efficient tool for genetic studies.

It has been suggested that SSR polymorphism results from two different mechanisms: slippage during replication and unequal crossing over (Li et al., 2002). In this study, we have found that the level of SSR polymorphism is positively correlated with the number of repeats or the length of SSR sequence (Table 3). These results suggest that most polymorphisms were generated by replication slippage (Innan et al., 1997) and that the number of repeats of a SSR is a useful predictor of its possible polymorphism (Yang et al., 1994). We have also found that SSRs with longer repeat motifs tend to be less polymorphic (Table 2). This phenomenon has been observed in *M. grisea* before (Wang et al., 2005) and also in human (Chakraborty et al., 1997). However, since opposite phenomenon was also observed (Schug et al., 1998; Ellegren, 2000), additional surveys are needed. In addition, we have also found that the polymorphism level of a SSR locus appears to be negatively correlated with the GC content of the SSR sequence (Table 2). This result implies that higher GC contents of SSR sequences could possibly reduce the probabilities of replication slippage and unequal crossing over.

In this study, we used the available physical map (version 5) as reference for constructing the genetic map. Most (except five) of the SSR markers with known positions in the physical map could

be assigned to expected chromosomes and reasonably linked following the orders in the physical map, although the parental strains used in this study (Guy11 and 2539) are different from the strain (70-15) used for the genome sequencing. The result suggests that the genome structures of different *M. grisea* strains are basically similar.

This similarity was further observed in the comparison between our map and the map reported by Wang et al. (2005) (denoted as W-map), which was constructed based on a cross (CH63 × TH16) different from that used in our study. There are 12 markers in our map that are equivalent to (i.e. based on the same SSR loci as) those from the W-map (Fig. 1 and Supplementary Table 1). The chromosomal allocations of these markers are all the same and the linkage orders of these markers are largely consistent between our map and the W-map.

We also compared our map with the maps of Nitta et al. (1997) (denoted as N-map) and Kaye et al. (2003) (K-map). The N-map mainly consists of RFLP markers and is based on the same cross used in our study (Guy11 × 2539). The K-map is an extension of the N-map with 23 SSR markers added. Eleven of these SSR markers have equivalent markers in our map (Fig. 1, Supplementary Table 1 and Supplementary Table 2). In addition, 72 RFLP markers in the N-map are expected to be closely linked with some SSR markers in our map according to the supercontigs and contigs (version 2) where the RFLP markers are located (http://www.broad.mit.edu/annotation/fungi/magnaporthe/markers.html#discrepancies; Supplementary Table 2). With these SSR and RFLP markers, our map and the K/N-map can be compared. Comparison indicates that our map and the K/N-map are largely consistent as expected (Supplementary Table 2). This validates the reliability of our map. Nevertheless, apparent inconsistencies are found for a few markers: five markers (33-8-H, CH3-24H, Pyrms93_94, G39R and CH2-59H) in the K/N-map are located far from the positions expected according to our map (e.g., 33-8-H is located at 21.2 cM from the first marker on chromosome 1 in the K/N-map, but the expected position according to our map should be between 3-9-H/64.9 cM and 1.2H/70.9 cM); five closely linked markers (55, 40, CH5-188H, 4-183 and G121R) on chromosome 4 in the K/N-map are located on chromosome 1 in our map (Supplementary Table 2). Since our map was constructed mainly according to the marker order in the physical map (version 5) of strain 70-15, it would be of interest to investigate whether the above discrepancies between the K/N-map and our map would reflect the genome difference between Guy11/2539 and 70-15. For this purpose, we constructed a genetic map without pre-assigning marker order (data not shown). Comparison indicates that the above discrepancies also exist between the K/N-map and this map. This result suggests that the discrepancies between the K/N-map and our map should not indicate putative genome differences between Guy11/2539 and 70-15. They might result from other causes. A possible cause is sampling/mapping errors. Another possible cause is that small genomic variation might exist between the parental strains we used and those used for the K/N-map although they should be very similar in principle. In addition, it is worth noting that the different chromosomal allocations of the region carrying the five markers 55, 40, CH5-188H, 4–183 and G121R in the K/N-map (on chromosome 4) and our map (on chromosome 1) happen to occur in parallel in the physical maps of version 2 (on chromosome 4) and version 5 (on chromosome 1). This coincidence of consistencies between the K/N-map and the physical map of version 2 and between our map and the physical map of version 5 and difference between the two groups of maps implies that translocation or duplication related to this region might exist between chromosome 1 and chromosome 4, if the discrepancy was not generated because of errors.

Magnaporthe grisea genome sequence (version 5) has at least 42 gaps and 266 contigs without known locations. Therefore, the genome sequence is not completed. In this study, 10 SSR markers from these contigs and 1 SSR marker developed from the gemome sequence of version 2 but not found in version 5 were mapped. We showed that they were mapped either to chromosome ends or to gaps in the physical map (Fig. 1). These results suggest that SSR mapping is helpful for filling gaps in the physical map although the reliability of the mapping result needs to be verified. It is noted that five of these markers without known locations in the physical map of version 5 were previously located in the physical map of version 2 on the same chromosomes as in our map, including MGM423 (on chromosome 1), MGM66 (chromosome 3), MGM388 (chromosome 4), MGM109 (chromosome 5) and MGM263 (chromosome 6). These results may give support to the locations of these five markers in our map and suggest that the previous chromosomal allocations of these markers (and corresponding contigs) in the physical map of version 2 might be correct. Moreover, a further validation was found for marker MGM109 on contig5.80, which is the same as ms419_420 from the W-map. The marker ms419_420 was located at one end of chromosome 5 in the W-map as MGM109 in our map. Therefore, contig5.80 is likely located at one end of chromosome 5 (Fig. 1).

Five markers located in the physical map (version 5) of 70-15 on chromosome 2 (MGM200, MGM201, MGM445 and MGM452) and chromosome 7 (MGM88) were mapped to chromosome 4 in our cross. Additionally, MGM201 was mapped far from MGM445 and MGM452 although they are located on the same contig (Fig. 1). This discrepancy is worth being further investigated. MGM452 is equivalent to marker Pyrms37_38 that is also mapped on chromosome 4 in the K-map. Therefore, this discrepancy between the physical map and our genetic map is not likely caused by sampling/mapping errors but reflect genetic differences between 70-15 and Guy11/2539.

It is known that centromeres greatly depress recombination rates in general (Davis et al., 1994; Farman and Leong, 1998). We have identified recombination cold spot(s) on each chromosome (Table 4). Most of these cold spots could be located in the K/W-map, although the numbers of markers used for these maps were not large enough for accurately determining these regions except two cold spots on chromosomes 3 and 6 with high densities of markers. This comparison suggests that these regions with low recombination rates are similar across different populations and experiments. Up to date, only the centromeric region of chromosome 7 has been genetically mapped in *M. grisea* (Thon et al., 2006). The cold spots identified in this study would provide clues for locating the centromeres of other chromosomes. However, recombination cold spots do not necessarily mean centromeric regions. According to Thon et al. (2006), the lowest recombination rate region on chromosome 7 of *M. grisea* is around *MAT* gene (this is consistent with our result), but the centromeric region of chromosome 7 is around marker MGM278 or Pyrms67_68 (Fig. 1).

Segregation distortion of molecular markers is a familiar phenomenon in genetic mapping and has been also reported in *M. grisea* (Romao and Hamer, 1992; Skinner et al., 1993; Farman and Leong, 1995; Nitta et al., 1997). In this study, we have found two large regions on chromosomes 2 and 6 showing significant segregation distortion in the map (Fig. 1). By comparing maps, we found that the segregation distortion region (MGM185–MGM47) on chromosome 2 contains a melanin biosynthesis gene *BUF1* (Farman, 2002). It was reported that melanin deficient mutation due to the deletion of *BUF1* gene took place frequently in the progeny of Guy11 × 2539 and most (87.5%) of the deletions occurred at the allele from Guy11 (Farman, 2002). In this study, 71 (23.7%) of the 299 progeny strains were melanin deficient. We randomly tested ten buff strains with markers MGM47 and MGM188 that flank *BUF1* gene and found that all of the buff strains inherited the alleles from Guy11. This explains why the segregation of this

region was biased to 2539, since the buff strains were all excluded from the mapping population. In regard to the segregation distortion region (MGM259–MGM402) on chromosome 6, which was biased to Guy11, the causes are not clear.

Acknowledgments

We are grateful to the three anonymous reviewers for their helpful comments and suggestions. This work was funded by the National Natural Science Foundation of China (Grant No. 30671123).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fgb.2008.07.012.

References

Adreit, H., Santoso, Andriantsimialona, D., Utami, D.W., Notteghem, J.L., Lebrun, M.H., Tharreau, D., 2007. Microsatellite markers for population studies of the rice blast fungus, *Magnaporthe grisea*. Mol. Ecol. Notes 4, 667–670.

Anderson, J.A., Churchill, G.A., Autrique, J.E., Sorrells, M.E., Tanksley, S.D., 1993. Optimizing parental selection for genetic linkage maps. Genome 36, 181–186.

Botstein, D., White, R.L., Skolnick, M., Davis, R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32, 314–331.

Brondani, C., Brondani, R.P.V., Garrido, L.R., Ferreira, M.E., 2000. Development of microsatellite markers for the genetic analysis of *Magnaporthe grisea*. Genet. Mol. Biol. 23, 753–762.

Budde, A.D., Smith, J.R., Farman, M.L., Skinner, D.Z., Leong, S.A., 1993. Genetic map and molecular karyotype of the blast fungus *Magnaporthe grisea*. In: O'Brien, S.J. (Ed.), Genetic Maps, sixth ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 3.110–3.111.

Chakraborty, R., Kimmel, M., Stivers, D.N., Davison, L.J., Deka, R., 1997. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proc. Natl. Acad. Sci. USA 94, 1041–1046.

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

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

Dean, R.A., 1997. Signal pathways and appressorium morphogenesis. Annu. Rev. Phytopathol. 35, 211–234.

Dean, R.A., Nicholas, J.T., Daniel, J.E., Farman, M.L., Mitchell, T.K., Orbach, M.J., Thon, M., Kulkarni, R., Xu, J.R., Pan, H., Read, N.D., Lee, Y.H., Carbone, I., Brown, D., Oh, Y.Y., Donofrio, N., Jeong, J.S., Soanes, D.M., Djonovic, S., Kolomiets, E., Rehmeyer, C., Li, W., Harding, M., Kim, S., Lebrun, M.H., Bohnert, H., Coughlan, S., Butler, J., Calvo, S., Ma, L.J., Nicol, R., Prucell, S., Nusbaum, C., Galagan, J.E., Birren, B.W., 2005. The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature 434, 980–986.

Ellegren, H., 2000. Microsatellite mutations in the germline: implications for evolutionary inference. Trends Genet. 16, 551–558.

Farman, M.L., 2002. Meiotic deletion at the *BUF1* locus of the fungus *Magnaporthe grisea* is controlled by interaction with the homologous chromosome. Genetics 160, 137–148.

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

Farman, M.L., Leong, S.A., 1998. Chromosome walking to the *AVR1-CO39* avirulence gene of *Magnaporthe grisea*: discrepancy between the physical and genetic maps. Genetics 150, 1049–1058.

Innan, H., Terauchi, R., Miyashita, N.T., 1997. Microsatellite polymorphism in natural populations of the wild plant *Arabidqsis thuliana*. Genetics 146, 1441–1452.

Karaoglu, H., Lee, C.M.Y., Meyer, W., 2004. Survey of simple sequence repeats in completed fungal genomes. Mol. Biol. Evol. 22, 639–649.

Kaye, C., Milazzo, J., Rozenfeld, S., Lebrun, M.H., Tharreau, D., 2003. The development of simple sequence repeat markers for *Magnaporthe grisea* and their integration into an established genetic linkage map. Fungal Genet. Biol. 40, 207–214.

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

Leung, H., Borromeo, E.S., Bernado, M.A., Notteghem, J.L., 1988. Genetic analysis of virulence in the rice blast fungus *Magnaporthe grisea*. Phytopathology 78, 1227–1233.

Li, Y.C., Korol, A.B., Fahima, T., Beiles, A., Nevo, E., 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol. Ecol. 11, 2453–2465.

Liu, R., Meng, J., 2003. MapDraw: a Microsoft Excel macro for drawing genetic linkage maps based on given genetic linkage data (in Chinese). Hereditas (Beijing) 25, 317–321.

Ma, J.H., Wang, L., Feng, S.J., Lin, F., Xiao, Y., Pan, Q.H., 2006. Identification and fine mapping of *AvrPi15*, a novel avirulence gene of *Magnaporthe grisea*. Theor. Appl. Genet. 113, 875–883.

Maroof, M.A.S., Biyashev, R.M., Yang, G.P., Zhang, Q., Allard, R.W., 1994. Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. Proc. Natl. Acad. Sci. USA 91, 5466–5470.

Martin, S.L., Blackmon, B.P., Rajagopalan, R., Houfek, T.D., Sceeles, R.G., Denn, S.O., Mitchell, T.K., Brown, D.E., Wing, R.A., Dean, R.A., 2002. MagnaprotheDB: a federated solution for integrating physical and genetic map data with BAC end derived sequence for the rice blast fungus *Magnaporthe grisea*. Nucleic Acids Res. 30, 121–124.

Nagakubo, T.M., Taga, M., Tsuda, M., Ueyama, A., 1983. Genetic linkage relationships in *Pyricularia oryzae* (in Japanese). Mere. Coll. Agric. Kyoto. Univ. 122, 75–83.

Nitta, N., Farman, M.L., Leong, S.A., 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., Hamer, J.E., 1992. Genetic organization of a repeated DNA sequence family in the rice blast fungus. Proc. Natl. Acad. Sci. USA 88, 5316–5320.

Rongwen, J., Akkaya, M.S., Bhagwat, A.A., Lavi, U., Gregan, P.B., 1995. The use of microsatellite DNA markers for soybean genotype identification. Theor. Appl. Genet. 90, 43–48.

Schug, M.D., Hutter, C.M., Wetterstrand, K.A., Gaudette, M.S., Mackay, T.F.C., Aquadro, C.F., 1998. The mutation rates of di-, tri- and tetranucleotide repeats in *Drosophila melanogaster*. Mol. Biol. Evol. 15, 1751–1760.

Schuler, G.D., 1997. Sequence mapping by electronic PCR. Genome Res. 7, 541–550.

Sesma, A., Osbourn, A.E., 2004. The rice blast pathogen undergoes developmental processes typical of root-infecting fungi. Nature 431, 582–586.

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

Skinner, D.Z., Leung, H., Leong, S.A., 1990. Genetic map of the blast fungus *Magnaporthe grisea*. In: O'Brien, S.J. (Ed.), Genetic Maps, fifth ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 3.82–3.83.

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

Struss, D., Plieske, J., 1998. The use of microsatellite markers for detection of genetic diversity in barley populations. Theor. Appl. Genet. 97, 308–315.

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

Sweigard, J.A., Valent, B., Orbach, M.J., Walter, A.M., Rafalski, A., Chumley, F.G., 1993. Genetic map of the rice blast fungus *Magnaporthe grisea*. In: O'Brien, S.J. (Ed.), Genetic Maps, sixth ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 3.112–3.115.

Talbot, N.J., Salch, Y.P., Margery, M.A., Hamer, J.E., 1993. Karyotypic variation within clonal lineages of the rice blast fungus, *Magnaporthe grisea*. Appl. Environ. Microbiol. 59, 585–593.

Tenzer, I., Ivanissevich, S.D., Morgante, M., Gessler, C., 1999. Identification of microsatellite markers and their application to population genetics of *Venturia inaequalis*. Phytopathology 89, 748–753.

Thon, M.R., Pan, H., Diener, S., Papalas, J., Taro, A., Mitchell, T.K., Dean, R.A., 2006. The role of transposable element clusters in genome evolution and loss of synteny in the rice blast fungus *Magnaporthe oryzae*. Genome Biol. 7, R16. doi: 10.1186/gb-2006-7-r16.

Tosa, Y., Osue, J., Eto, Y., Oh, H.S., Nakayashiki, H., Mayama, S., Leong, S.A., 2005. Evolution of an avirulece gene, *AVRI-CO39*, concomitant with the evolution and differentiation of *Magnaporthe oryzae*. Mol. Plant Microbe Interact. 18, 1148–1160.

Valent, B., 1990. Rice blast as a model system for plant pathology. Phytopathology 80, 33–36.

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

Wang, B.H., Lu, G.D., Li, H.M., Lin, Y., Wang, Z.H., 2004. Assessment of *Magnaporthe grisea* mating type by PCR (in Chinese). J. Agr. Biotechnol. 12, 183–187.

Wang, Y., Kaye, C., Bordat, A., Adreit, H., Milazzo, J., Zheng, X., Shen, Y., Thearreau, D., 2005. Construction of genetic linkage map and location of avirulence genes from cross CH63 and TH16 of *Magnaporthe grisea* (in Chinese). Chinese J. Rice Sci. 19, 160–166.

Xu, S., Tao, Y., Yang, Z., Chu, J., 2002. A simple and rapid methods used for silver staining and gel preservation (in Chinese). Hereditas (Beijing) 24, 335–336.

Yang, G.P., Maroof, M.A.S., Xu, C.G., Zhang, Q., Biyashev, R.M., 1994. Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. Mol. General Genet. 245, 187–194.