POLYMORPHISM OF MICROSATELLITE SEQUENCE WITHIN PROTEIN KINASE ORFS IN PHYTOPATHOGENIC FUNGUS, *MAGNAPORTHE GRISEA*

Chengyun Li[1,*], Lin Liu[1], Jing Yang[1], Jinbin Li[2], Zhang Yue[1], Yunyue Wang[1], Yong Xie[1], Youyong Zhu[1]

[1] *Key Laboratory for Agricultural Biodiversity and Pest Management of China Education Ministry, Plant Protection College, Yunnan Agricultural University, Kunming, 650201, China*

[2] *Plant Protection Research Institute of Yunnan Academy of Agricultural Sciences, Kunming, 650205, China*

* *Correspondence: Chengyun Li, Key Laboratory for Agricultural Biodiversity and Pest Management of China Education Ministry, Yunnan Agricultural University, Hei Longtan, Kunming, 650201, China, Fax: 86-871-5227945, Email: li.chengyun@gmail.com*

Abstract: Eighteen polymorphic microsatellite markers suitable for population genetic studies and protein kinase encoding genic variation measurement were developed for rice blast fungus *Magnaporthe grisea*. Polymorphism was evaluated by using 46 isolates collected from diverse geographical locations and rice varieties. Preliminary results indicate that each locus harbors two to fourteen alleles.

Keywords: *Magnaporthe grisea*; protein kinase; microsatellite

1. INTRODUCTION

Magnaporthe grisea is the most destructive pathogen of rice worldwide and the primary model organism for elucidating the molecular basis of fungal diseases of plants (Valent, 1990) The completion of the genome

Li, C., Liu, L., Yang, J., Li, J., Yue, Z., Wang, Y., Xie, Y. and Zhu, Y., 2008, in IFIP International Federation for Information Processing, Volume 258; Computer and Computing Technologies in Agriculture, Vol. 1; Daoliang Li; (Boston: Springer), pp. 559–563.

sequence for *Magnaporthe grisea* has made it possible to determine the total number of genes as well as to analyze and classify them according to their structure and function (Dean et al., 2005).

The eukaryotic protein kinases comprise one of the largest superfamilies of homologous proteins and genes. There are now hundreds of different members whose sequences are known within this family (Hanks, 2003). Although there are common structural features among protein kinases, differences in structural features, regulation modes, and substrate specificities divide them into separate groups. In the phytopathogenic fungi, components of heterotrimeric G proteins, MAP kinases, and cAMP signal transduction pathway are required for pathogenesis (Muller et al., 2003; Yamada-Okabe et al., 1999; Xu et al., 1996). We previous revealed that many protein kinase genes harbored SSR sequences within their protein coding region (Li C.Y. et al., 2005), but whether these sequences are polymorphic is unclear.

Microsatellites are favored for genetic applications because they are abundant in plant genomes, highly polymorphic within species, relatively rapid and inexpensive to assay, and can be used to identify specific chromosomal regions consistently across populations. Distribution and frequency of SSRs in genomic scale or ESTs have been analyzed extensively, however, reports published to date clearly discussed SSR polymorphism in genes has been limited (Li C.Y. et al., 2005).

The objective of this study was to develop PCR primer pairs targeting previously sequenced genes from *M. grisea* in order to compare the allelic amplification product polymorphism of protein kinase encoding genes among natural populations.

2. MATERIAL AND METHOD

The DNA sequence and a database of known and predicted open reading frames of eukaryotic protein kinases were obtained from the *Magnaporthe grisea* genome database (http://www.genome.wi.mit.edu/annotation/fungi/magnaporthe/) on July 14, 2005. The program software tandem repeats finder (TRF) written by Benson (Benson, 1999) with the following options: minimum size = 15 bp, 80% matches (namely number of matched bases between two repetitive elements is 80%) and abundance were removed.

Polymorphic loci were detected by screening a subset of 46 *M. grisea* isolates collected from different locations and from various rice varieties of Yunnan Province, China. Genomic DNA was extracted from mycelia using a

simple extraction protocol (Zhang et al., 1996). Primers were designed for DNA sequence with microsatellite motifs using PRIMER3 (Rozen et al., 2000) software and synthesized by Invitrongen Biotechnology Co. Ltd. Shanghai, China.

PCR amplifications were carried out in 20 μL volumes containing 1 × PCR buffer (10 mM Tris-HCl pH 8.5, 50 mM KCl, 1.5 mM $MgCl_2$, and 0.001% gelatin), 125 μM each dNTP, 5 pmol of each primer, and 0.5 U of *Taq* DNA polymerase (Sino-American Biotechnology Co., Beijing). Approximately 50 ng of genomic DNA was used for each reaction. Amplification were performed in a Eppendoff PCR thermal Matercycler with the cycling parameters; 5 min at 94°C, 35 cycles of 1 min at 94°C, 1min at 54°C and 1 min at 72°C followed by a final extension for 10 min at 72°C. In initial experiments, amplified fragments were visualized by electrophoreses in 1.5% agarose gels stained with ethidium bromide. Loci that appeared polymorphic were further examined by 8% polyacrylamide gel to determine the product size of the PCR product and number of alleles per locus. Fragment size of PCR products were estimated on Bio-Imaging System E5000.

3. RESULTS

Eighteen of the 26 polymorphic loci produced amplicons from a majority of 46 isolates, and displayed anywhere from two to fourteen alleles (Table 1). Gene diversity was estimated with the software program, GENEPOP (V1.32), and are shown in table 1. KMS02, KMS07, KMS18 showed high gene diversity in population used for the study. This suggests that genes harboring these SSR sequences are also highly diverse within the populations.

M. grisea has 11 109 proteins coding ORFs in whole genome, and among these, 85 protein kinase genes, corresponding to – 0.76% of the total number of genes (Dean et al., 2005). More than 30% of these protein kinase encoding genes have SSRs within their protein coding regions. The high degree of polymorphism in this set of microsatellite markers can be used to analyze population structure and strain distribution in association with rice variety and location, adding to the fundamental understanding the function of protein kinase genes of the fungus. These results provide useful information to study possible SSR functions and variation of protein kinases that harbor them.

Table 1. Polymorphsims of SSRs in protein kinase encoding ORFs in *M. grisea.* N_a, number of alleles; G_d: Gene diversity by Shannon's Information inde

Locus	Contig*	Primer (F,5'-3')	Primer (R,5'-3')	Gene name	Motif	No of Genotypes	Repeat No.	Na	Percent matches	Product size range (bp)	Gd
KMS012	1190	AGCGAAACAA GAACGCGAGG	TAGCAGCCTG TGCTCGTTCG	MGG 06413.5	ACG	46	5	2	100	219-222	0.5623
KMS022	1299	CGCAAAGAAT TCAAACCCGC	GGAGACGACA CTGGGGTGCT	MGG 07003.5	ACG	46	6	14	100	279-326	2.4256
KMS032	1302	TCCCTTTCGGT CGTTCCAAG	CCCGCTGAGG TAGCCAAAGA	MGG 07012.5	ACG	46	5	4	86	208-220	1.0673
KMS042	1631	CCGAAGAGGT CCTCCAAGCA	GCAGCAGCAT TACCATCCCC	MGG 08643.5	AG	46	6	5	88	216-311	0.654
KMS052	377	CCAAGCCCAG AGCCAGAAAA	CTCGAGGCTG CCCATCTTGT	MGG 01998.5	AGA	46	5	3	85	243-249	0.2826
KMS062	1692	GCATGAAATG CTCGTCGTGG	ATGCAGCGGC CTTCATTAGC	MGG 09000.5	AGAAAA	46	3	4	100	191-209	1.181
KMS072	347	GTTCTCCATCG CCCAAATCG	TAAGTGGGCT CCTTGGCTGC	MGG 01816.5	CAA	46	5	10	86	245-281	2.1018
KMS092	993	GGTCCTATTCC CCTTCCCCC	TTCAACCGAT ACGAGGCCGT	MGG 05397.5	CAAGG	46	3	2	90	200-206	0.5921
KMS112	211	GGGGTAACGA CAGCCAGGTG	CTGCTGCTGCT GTTGCTGGT	MGG 01196.5	CAG	46	7	5	80	218-230	1.1916
KMS122	1645	CGATACCCCT GAGCCACCAC	ATCTGCCGCCT TTTGAGTGC	MGG087 46.5	CAG	46	5	6	100	198-213	1.3926
KMS132	1353	GTCATGACAG GGGTCCTCGG	ATCCCGCTCTC CCTCCATTC	MGG 07291.5	CAGCCC	46	3	8	84	232-292	1.8262
KMS142	1182	GAACCCCGCA GTCCAACAAC	CGCCTGCTGA GAATGGGACT	MGG 06368.5	CAGTCA	46	3	5	92	205-229	1.2688
KMS152	838	GGCCACAGAG GAGAACGGAA	TCCATCAGGA TCGGGGACTG	MGG 04463.5	CGA	46	5	6	83	215-233	1.7452
KMS162	343	TCATGAGCGA GACAATGGGG	TGCTGAACCG ATTCCGCTTT	MGG 01795.5	CGGT	46	4	2	83	206-210	0.5623
KMS182	687	GCAAGTCGCC TCGCCATTAT	TGAACCGACT CGTCGACTGC	MGG 03488.5	GCA	46	6	5	82	244-256	1.4555
KMS192	1360	CAGCACCCAA AAGGAGCCTG	AACATTCCCA GGTGCATCGC	MGG 12406.5	GCA	46	5	2	85	218-221	0.3594
KMS202	1779	CGCCCTTCAA AAACCAAGGG	AATTGCGACA AGTCGCTCCC	MGG 14499.5	CAG	46	7	5	100	257-281	1.3841
KMS222	338	AGACGACGAG GCTTCCGATG	GGCATAAGGT TGTCGCGGAG	MGG 00925.5	GCT	46	6	5	100	187-211	0.8856

* Contig is based on the *M. griseea* genome database,
website: http://www.genome.wi.mit.edu/annotation/fungi/magnaporthg/

ACKNOWLEDGEMENTS

We thank Miss J. Krenz, in Department of Botany and Plant Pathology, Oregon State University of USA for helpful advice and revision the paper in

detail. This work is supported by National Basic Research Program of China (2006CB100202), Education Ministry Foundation (307025) and Doctorial Foundation of Education Ministry of China (20050676001).

REFERENCES

Benson G. 1999, Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res., 27: 573-580.

Dean R.A., Talbot N.J., Ebbole D.J. et al. 2005, The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature, 434: 980-986.

Hanks S.K. 2003, Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. Genome Biol., 4, 111.

Li C.Y, Li J.B., Zhou X.G., Zhang S.S., Dong A.R., Xu M.H. 2005, Frequency and distribution of microsatellites in open reading frame of rice blast fungus, *Magnaporthe grisea*. Chinese J. Rice Sci., 19: 167-173. 2005. (in Chinese with English abstract)

Muller P., Weinzierl G., Brachmann A., Feldbrugge M., Kahmann R. 2003, Mating and pathogenic development of the Smut fungus Ustilago maydis are regulated by one mitogen-activated protein kinase cascade. Eukaryot. Cell, 2: 1187-1199.

Rozen S., Skaletsky H. 2000, Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol., 132: 365-386.

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

Xu J.R., Hamer J.E. 1996, MAP kinase and cAMP signaling regulate infection structure formation and pathogenic growth in the rice blast fungus *Magnaporthe grisea*. Genes Dev., 10: 2696-2706.

Yamada-Okabe T., Mio T., Ono N., Kashima Y., Matsui M., Arisawa M., Yamada-Okabe H. 1999, Roles of three histidine kinase genes in hyphal development and virulence of the pathogenic fungus Candida albicans. J. Bacteriol., 181: 7243-7247.

Zhang D., Yang Y., Castlebury L.A., Cerniglia C.E. 1996, A method for the large scale isolation of high transformation efficiency fungal genomic DNA. FEMS Microbiol. Lett., 145: 261-265.