Trop. plant pathol. (2016) 41:331–335
DOI 10.1007/s40858-016-0104-6

SHORT COMMUNICATION

Mapping highly informative SSR markers in the genome of *Magnaporthe oryzae* from wheat

Antonio Nhani Júnior[1] · **Jorge Fernando Pereira**[1,2] · **Jéssica Rosset Ferreira**[3] · **Ana Lídia Variani Bonato**[1] · **João Leodato Nunes Maciel**[1]

Received: 13 June 2016 / Accepted: 7 September 2016 / Published online: 26 September 2016

Abstract Wheat blast, caused by *Magnaporthe oryzae* (*Triticum* haplotype - MoT), is an important disease of wheat in Brazil. In this study, we designed 38 new SSR markers based on the genome sequence of different MoT isolates, compared the informativeness of those markers with other 52 from the literature and mapped the polymorphic ones. Among the 90 SSR markers, 53 were polymorphic resulting in, on average, 3.02 alleles per locus and polymorphism information content (PIC) of 0.41. Most (81.1 %) of the polymorphic markers presented 11 or more motif repeats. Seventeen highly informative markers were detected and mapped in all chromosomes except for chromosome 5. On average, polymorphic markers on chromosome 6 showed the highest PIC followed by chromosomes 2 and 7. Clustering analysis showed a clear separation of one isolate from rice from the rest of the isolates from wheat. In each of the three clusters detected, the MoT isolates were similar among them regardless of the year and location sampled, suggesting that the pathogen is widely dispersed across wheat growing regions in Brazil. The highly informative markers detailed here should be useful for population biology studies of the wheat blast pathogen.

Keywords Genetic variability · Microsatellite · Polymorphism information content · Wheat blast

Magnaporthe oryzae B.C. Couch and L.M. Kohn (anamorph. *Pyricularia oryzae* Cavara) is the cause of the blast disease in many hosts, but strains causing wheat blast have been grouped in the *Triticum* haplotype (MoT). *Magnaporthe oryzae* is an heterothallic ascomycete whose dispersion occurs via air currents and infected seeds (Urashima et al. 2004, 2007). The disease is spread over all wheat-growing regions in Brazil, and some of the neighbor countries in South America. Recently, it has been detected in Bangladesh (Callaway 2016).

The severity of wheat blast is influenced by the genetic variability of the pathogen since, in field conditions, the tolerance of a specific wheat genotype is not always confirmed in different geographical regions (Urashima et al. 2004). Moreover, although there are wheat lineages and cultivars showing resistance to one or few MoT isolates, no resistant wheat genotype, including synthetic wheat, was observed when a higher number of MoT isolates was tested (Cruz et al. 2009). In this context, it is imperative to understand the genetic variability, population structure and migration patterns of the MoT isolates in order to improve phenotyping methods in controlled conditions. These methods could use different MoT isolates whose genetic variability is representative of the one detected in field conditions. To achieve that, reliable molecular methods are required. Nowadays, genome sequencing is a viable alternative to evaluate genetic variability. It can also answer questions about the emergence of wheat blast, genes involved in host specificity, among others (Thynne et al. 2015). However, the easy detection and multiallelic nature

Antonio Nhani Júnior and Jorge Fernando Pereira contributed equally to this work.

Section Editor: Robert N.G. Miller

Electronic supplementary material The online version of this article (doi:10.1007/s40858-016-0104-6) contains supplementary material, which is available to authorized users.

✉ João Leodato Nunes Maciel
joao.nunes-maciel@embrapa.br

[1] Embrapa Trigo, CEP 99050-970 Passo Fundo, RS, Brazil

[2] Present address: Embrapa Gado de Leite, CEP 36038-330 Juiz de Fora, MG, Brazil

[3] Universidade de Passo Fundo, CEP 99052-900 Passo Fundo, RS, Brazil

of SSR markers are useful for a number of studies where genome sequencing and bioinformatics analysis are still a not trivial endeavor. Our group has routinely used SSR markers to study MoT isolates although, even when using SSR markers described by others (Cruz et al. 2009) or obtaining new ones from an enriched microsatellite library (Pereira et al. 2014), most of the SSR markers were monomorphic or slightly informative (PIC < 0.25).

In this study, we used genomic sequences of nine MoT isolates (A. Nhani Jr - unpublished results) to design 38 new SSR markers. We compared these markers with other 52 previously described by evaluating all the 90 markers in the same set of isolates. Our objective was to evaluate the informativeness of all the 90 SSR markers and to map the highly informative ones in the genome of *M. oryzae* isolated from wheat. Additionally, we also evaluated the genetic variability of 27 MoT isolates from different wheat areas of seven Brazilian states.

The 27 strains of *M. oryzae* isolated from wheat, which were collected in seven Brazilian states, and one isolated from rice used in this work are listed in Table 1. The maintenance of the isolates and the DNA extraction were performed as described by Pereira et al. (2014). The SSR markers were detected in the sequenced genome of nine MoT isolates (Table 1) and in the reference genome *Magnaporthe oryzae* 70-15 (Dean et al. 2005) version 8.0, obtained at the Magnaporthe Comparative Genomic Database (http://www.broadinstitute.org/annotation/genome/magnaporthe_comparative/MultiDownloads.html). The perl script MISA (MIcroSAtellite; http://pgrc.ipk-gatersleben.de/misa/) was configured to search for perfect motifs (zero interruptions) and di- to hexanucleotides with a minimum of 11 repetitions. *In house* perl scripts were used for screening of polymorphic SSR loci through comparison of region sizes and positions among all MISA results. Primers were designed using the Primer3 software (Rozen and Skaletsky 2000). The three-primer system (Schuelke 2000) was used where a M13 tail (TGTAAAACGACGGCCAGT) was added to the forward primers and, along with the reverse primers and an M13 primer labeled with a fluorescent dye (FAM, NED, PET, or VIC), used for the amplifications. Each SSR marker was amplified using 25 ng of total DNA and a common reaction mixture containing 1× buffer, 2.5 mM $MgCl_2$, 0.2 μM of reverse primer, 0.02 μM of forward primer, 0.2 μM of labeled M13 primer and 0.75 U of *Taq* polymerase. Two concentrations of each dNTP were used: 0.2 mM (mix 1) or 0.35 mM (mix 2). For each SSR primer, only one reaction mixture was chosen (Supplementary Table 1). Amplification was performed in a GeneAmp® PCR System 9700 (Applied Biosystems) and, depending on the melting temperatures of the primers, programs TD60-50 or TD60-55 were used. The program TD60-50 consisted of ten cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s, where 1 °C was decreased per cycle and then 25 cycles at 94 °C for 30 s, 50 °C for 30 s, 72 °C for 30 s. The program TD60-55 consisted of five cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s, where 1 °C was decreased per cycle and then 30 cycles at 94 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s. Both programs had one denaturation step at 94 °C for 3 min before the amplification cycles and one amplification step of 72 °C for 15 min at the end. After amplification, reactions were run on an ABI 3130xl Sequence Analyzer containing a 36 cm capillary array with POP6 polymer after being diluted in water, mixed with Hi-Di formamide and GeneScan 500 LIZ size standard (Applied Biosystems) and denatured (95 °C for 5 min). The program GeneMapper v4.1 was used to analyze the data (Supplementary Fig. 1).

The number of alleles, average of alleles per locus and polymorphism information content (PIC) were calculated using the PowerMaker program (Liu and Muse 2005). The polymorphism information content was estimated according to Anderson et al. (1993): $PIC = 1 - \sum P_{ij}^2$, where P_{ij} is the frequency of the *i*th allele at the *j*th marker, to evaluate the diversity level of each SSR marker. PowerMarker was also used to investigate the genetic similarity between accessions. Therefore, an Unweighted Pair Group Method with Arithmetic Mean (UPGMA) tree was constructed using the shared allele distance for each pair of individuals. The tree was based on the matrix of shared microsatellite alleles among the 28 accessions calculated from the 90 SSR markers. All the accessions and markers were also used for population structure analysis performed in Structure 2.3 software (Pritchard et al. 2000). This program implements a Bayesian clustering procedure to assign individuals to clusters (*K*) without prior knowledge of their population affinities. The likelihood of *K* is estimated from the allele frequencies. Structure was run with 10 independent simulations, each one using 50,000 burn-in periods, following 100,000 MCMC repetitions. The best-supported value of *K* was identified using Structure Harvester (Earl 2012) and the results were plotted using Pophelper (Francis 2016). Besides that, the polymorphic markers detected here were anchored to the genome of the isolate Py 5003 using Ugene (Okonechnikov et al. 2012) embedded Burrows-Wheeler Aligner (Li and Durbin 2009) with options seed length 18 and max diff zero.

Among the 38 new SSR markers (MoO_01 to MoO_58) evaluated in this study, 12 (31.5 %) amplified a monomorphic fragment when the rice isolate (I.162) was not considered. This is lower than the 60.5 % monomorphic SSR loci previously characterized from one MoT isolate (Pereira et al. 2014). One explanation for the decrease in monomorphic SSR loci is the use of at least 11 motif repeats among the di-, tri- or tetranucleotides markers in comparison with the minimum of three repeats used by Pereira et al. (2014). In fact, when considering all the 90 SSR markers evaluated here, only 10 out of the 53 polymorphic markers were based on 10 or lower motif repeats (Supplementary Table 1). In this way, when

Table 1 Information for the collection of *Magnaporthe oryzae* isolates used in this study.

Identification[a]	Municipality/State of origin[b]	Year of isolation
I.162	São Miguel do Araguaia / GO	2008
Py 6029 and Py 6045*	Goiânia / GO	2006
Py 12.1.050 and Py 12.1.062	Rio Verde / GO	2012
Py 10121.1	Brasília / DF	2010
Py 12.1.109, Py 12.1.110, Py 12.1.116 and Py 12.1.154	Brasília / DF	2012
Py 30.2 and Py 35.3	Brasília / DF	2008
Py 36.1*	Brasília / DF	2007
Py 12.1.170	Amambai / MS	2012
Py 12.1.321 and Py 12.1.323	Aral Moreira / MS	2012
Py 6007, Py 6017* and Py 6025*	Coromandel / MG	2006
Py 12.1.094	Perdizes / MG	2012
Py 0925*	Perdizes / MG	2009
Py 12.1.035i, Py 12.1.046i and Py 12.1.051i	Itaí / SP	2012
Py 5003*, Py 5010* and Py 5033*	Londrina / PR	2005
Py 5007	Londrina / PR	2006
Py 11.118.3	Londrina / PR	2011
Py 12.1.017i and Py 12.1.328	Londrina / PR	2012
Py 12.1.298	Rolândia / PR	2012
Py 86.1*	Cascavel / PR	2008
Py 12.1.179 and Py 12.1.188	São Luís Gonzaga / RS	2012
Py 12.1.209	São Borja / RS	2012

[a] I.162 was isolated from rice and all other isolates were obtained from infected wheat plants found in farmer fields. Asterisks represent the nine isolates whose genomes were used to detect the 38 new SSR Markers described in this study. Among these isolates, only Py 6017 was used for the clustering analysis. [b] DF means Distrito Federal while GO, MS, MG. SP, PR and RS indicate the Brazilian states of Goiás, Mato Grosso do Sul, Minas Gerais, São Paulo, Paraná and Rio Grande do Sul, respectively.

evaluating the variability of MoT isolates through SSR markers, the use of loci showing a higher number of repeat motifs is greatly recommended.

Considering all the 90 SSR markers, we detected 53 polymorphic loci among the MoT isolates with an average of 3.02 alleles per locus. The PIC value of each polymorphic marker varied from 0.07 to 0.72 with an average of 0.41. The average number of alleles and PIC obtained here were higher than previously reported for MoT isolates (Cruz et al. 2009; Pereira et al. 2014) but lower than studies with *M. oryzae* from rice (Adreit et al. 2007; Zheng et al. 2008). In order to rank the markers' informativeness, we separate them in highly (PIC > 0.50), reasonably (PIC from 0.25 to 0.50) and slightly informative (PIC < 0.25), as proposed by Botstein et al. (1980). Highly informative markers represented 32 % (17 loci) of the polymorphic ones being remarkably different from one study with *M. oryzae* isolate from rice where 53.7 % of the SSR markers were found to be highly informative (Zheng et al. 2008). Among the 17 highly informative SSR markers, 10 were designed here, six were also developed from one MoT isolate (Pereira et al. 2014) and only one from a rice isolate (Garrido 2001). In fact, the markers MGM428 and MGM429 were among the most polymorphic ones described for *M. oryzae* isolated from rice (Zheng et al. 2008) but here were only slightly informative with PIC = 0.20 for both markers (Supplementary Table 1). Our mapping, using the genome of one MoT isolate named Py 5003 (A. Nhani Jr - unpublished results), revealed that all chromosomes, except for the chromosome 5, showed markers with PIC > 0.50 (Fig. 1). The average of PIC values was higher for markers on chromosome 6 followed by chromosomes 2 and 7 and lower for chromosome 5. For two markers (Pyrms87-88 and MGM-1) only the reverse or forward primer could be mapped. Moreover, the marker Pyrms87-88 from chromosome 1 (Kaye et al. 2003) was mapped to chromosome 2 and markers MGM428 and MGM429 from chromosome 5 (Zheng et al. 2008) were mapped on chromosome 4. One region with low variability (five markers with PIC < 0.25) was found on chromosome 4.

The SSR alleles detected here allowed the construction of a dendrogram where the rice isolate (I.162) was clearly separated from the MoT isolates (Fig. 2). Our analysis also showed similarity among isolates from distant locations. For instance, the isolates from RS (the southernmost State in Brazil) are similar to one isolate from the PR State or one from the MS State, meaning that the isolates established in RS were

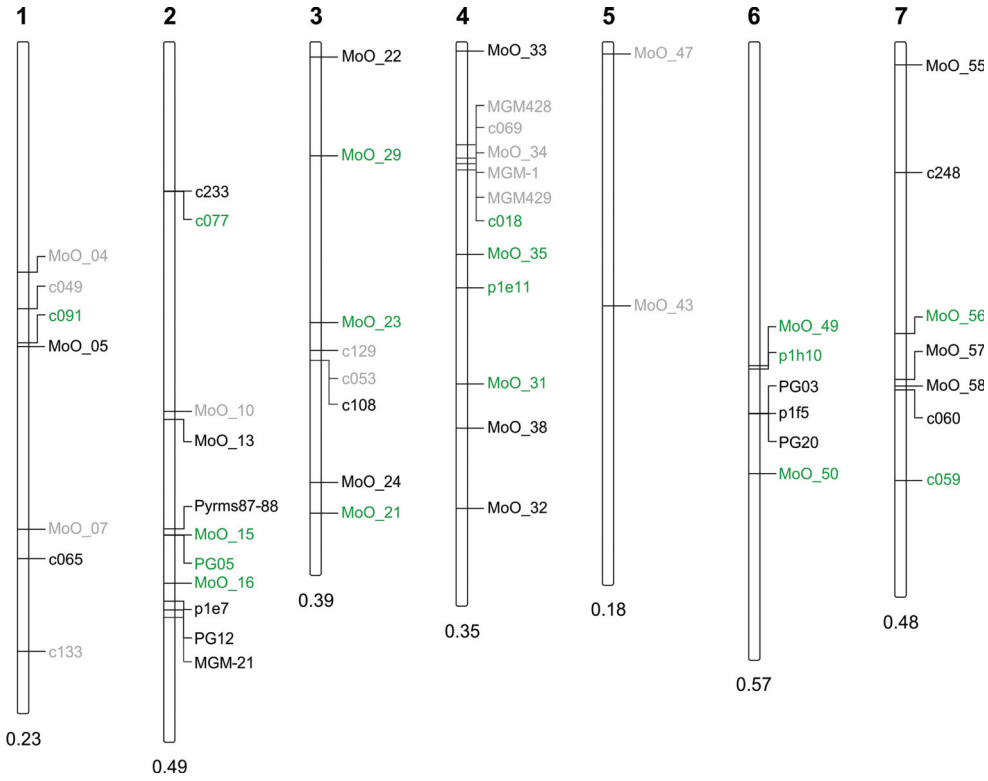

Fig. 1 Map of the 53 polymorphic SSR markers in the genome of the MoT isolate Py5003. The slightly informative markers (PIC < 0.25) are represented by *gray letters*, while the reasonably informative (0.25 < PIC < 0.50) are in *black* and the highly informative (PIC > 0.50) ones are in *green*. Numbers bellow each of the seven chromosomes represent the average PIC of the markers located in that chromosome

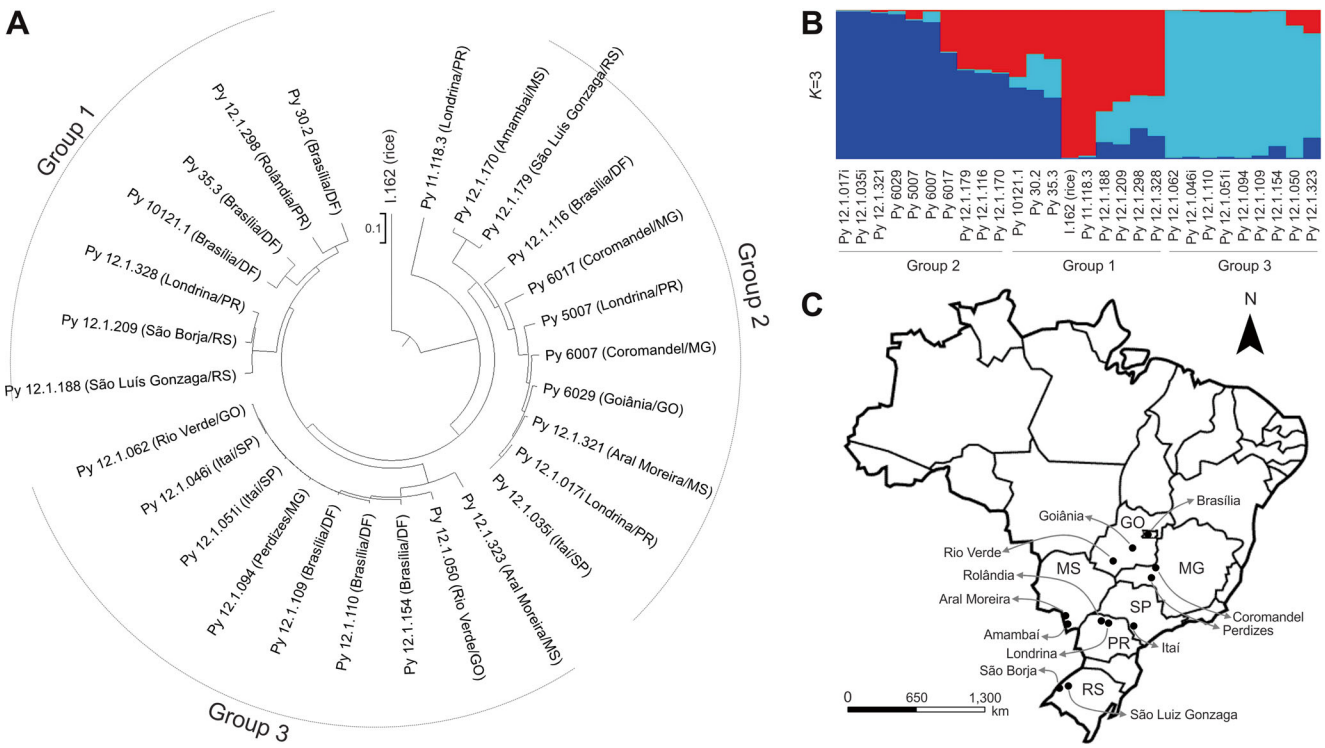

Fig. 2 Analysis of 28 *M. oryzae* isolates with the 90 SSR markers evaluated in this study. **a** Clustering dendrogram based on the alleles detected in one rice isolate (I.162) and 27 wheat isolates. **b** Estimated proportion of membership in the corresponding clusters ($K = 3$) as calculated using Structure software. **c** Locations where the isolates were collected. *Dots* represent 12 cities located in the States of Goiás (GO), Mato Grosso do Sul (MS), Minas Gerais (MG), Paraná (PR), São Paulo (SP), and Rio Grande do Sul (RS). Brasília is located in the Distrito Federal

possibly originated from different regions. Moreover, isolates from the same location (for example, Brasília/DF, Londrina/PR and Itaí/SP) were distributed across different clusters and, interestingly, Group 2 is formed by isolates from all the seven States analyzed. That reveals movement of the pathogen across the country resulting in genetic similarity even for isolates from wheat fields separated for around 2,000 km. Infected wheat seeds transported across different regions might be an explanation for that (Urashima et al. 2004). The population structure analysis, where the best inferred number of clusters obtained was $K = 3$, was in agreement with the UPGMA results (Fig. 2B).

In conclusion, when comparing to previous studies, we detected a greater number of highly informative SSR markers for *M. oryzae* isolated from wheat. Moreover, considering all polymorphic markers, we have obtained a higher average number of alleles per locus and higher average PIC. The reason for that probably relies on the use of SSR loci showing higher number of repeat motifs (≥11). The highly informative SSR markers were mapped in the genome of the MoT isolate Py 5003 revealing that the average PIC values of the markers on chromosomes 2, 6 and 7 was higher in contrast with low values for chromosomes 1 and 5. Similarity among isolates from distant geographic regions was detected indicating a constant movement of the pathogen in different wheat growing areas in Brazil. The highly informative SSR markers identified here should be useful for molecular studies of *M. oryzae* from wheat.

Acknowledgments We are thankful to Embrapa (Empresa Brasileira de Pesquisa Agropecuária) for the financial support (Project 02.11.04.006.00.00 - Variação genética e de virulência de *Magnaporthe oryzae* do trigo e de Poáceas invasoras) and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for the PhD scholarship of JRF. We thank Jordalan Buffet Muniz for assistance with DNA extraction and quantification. ANJ thanks Professor Nicholas J. Talbot for the research support.

References

Adreit H, Andriantsimialona DS, Utami W, Notteghem JL, Lebrun MH, Tharreau D (2007) Microsatellite markers for population studies of the rice blast fungus, *Magnaporthe grisea*. Mol Ecol Notes 7:667–670

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36:181–186

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Callaway E (2016) Devastating wheat fungus appears in Asia for first time. Nature 532:421–422

Cruz MF, Maciel JLN, Prestes AM, Bombonatto EAS, Pereira JF, Consoli L (2009) Caracterização genética e fenotípica de isolados de *Pyricularia grisea* do trigo. Trop Plant Pathol 34:393–401

Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan H et al (2005) The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature 434:980–986

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

Francis RM (2016) POPHELPER: an R package and web app to analyse and visualize population structure. Mol Ecol Resour. doi:10.1111/1755-0998.12509

Garrido LR (2001) Desenvolvimento e uso de marcadores de regiões hipervariáveis do genoma de *Magnaporthe grisea* na análise da estrutura de populações do patógeno infectando plantações de arroz. Tese de Doutorado. Universidade de Brasília. Brasília, DF

Kaye C, Milazzoa J, Rozenfeldb S, Lebrunc MH, 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

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760

Liu K, Muse SV (2005) PowerMarker: integrated analysis environment for genetic marker data. Bioinformatics 21:2128–2129

Okonechnikov K, Golosova O, Fursov M (2012) Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28:1166–1167

Pereira JF, Consoli L, Bombonatto EAS, Bonato ALV, Maciel JLN (2014) Development of genomic SSR markers and molecular characterization of *Magnaporthe oryzae* isolates from wheat in Brazil. Biochem Genet 52:52–70

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, pp 365–386

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234

Thynne E, McDonald MC, Solomon PS (2015) Phytopathogen emergence in the genomics era. Trends Plant Sci 20:246–255

Urashima AS, Lavorent NA, Goulart ACP, Mehta YR (2004) Resistance spectra of wheat cultivars and virulence diversity of *Magnaporthe grisea* isolates in Brazil. Fitopatol Bras 29:511–518

Urashima AS, Leite SF, Galbieri R (2007) Eficiência da disseminação aérea em *Pyricularia grisea*. Summa Phytopathol 33:275–279

Zheng Y, Zhang G, Lin F, Wang Z, Jin G, Yang L, Wang Y, Chen X, Xu Z, Zhao X, Wang H, Lu J, Lu G, Wu W (2008) Development of microsatellite markers and construction of genetic map in rice blast pathogen *Magnaporthe grisea*. Fungal Genet Biol 45:1340–1347