2011

Discriminating microsatellites from *Macrophomina phaseolina* and their potential association to biological functions

R. S. Arias
USDA-ARS, renee.arias@ars.usda.gov

J. D. Ray
USDA-ARS

A. Mengistu
USDA-ARS

B. E. Scheffler
USDA-ARS

Plant Pathology (2011) Doi: 10.1111/j.1365-3059.2010.02421.x

Discriminating microsatellites from *Macrophomina phaseolina* and their potential association to biological functions

R. S. Arias[a]*, J. D. Ray[b], A. Mengistu[c] and B. E. Scheffler[a]

[a]*USDA-ARS MSA Genomics and Bioinformatics Research Unit, 141 Experiment Station Road, Stoneville, MS 38776;* [b]*USDA-ARS Crop Genetics Research Unit, 141 Experiment Station Road, Stoneville, MS 38776; and* [c]*USDA-ARS Crop Genetics Research Unit, 605 Airways Boulevard, Jackson, TN 38301, USA*

One hundred and eighty-two microsatellites or simple-sequence-repeat (SSR) markers for *Macrophomina phaseolina* were developed. These were tested on 24 isolates of *M. phaseolina* obtained from seven plant species, and the genetic variation of isolates was studied in relation to potential biological processes that could be affected in this fungus. A total of 120 SSR markers were polymorphic, amplifying >90% of the 24 isolates tested. Thirty percent of the markers showed multiple alleles on individual samples. A large number of markers showed unique alleles in isolates collected from pumpkin and snap bean. DNA sequences corresponding to 43 markers had significant hits on BLASTX and/or BLAST2GO, and the polymorphism of 36 of those markers showed specific allele patterns for one or more plant host origin of the isolates. Additional tests on growth rate and copper resistance of the isolates identified markers that could be related to those traits. In addition, 27 markers were monomorphic and amplified all 24 isolates. Whereas polymorphic markers can be used for population genetics studies of *M. phaseolina*, the group of 27 monomorphic markers could help in the fast identification of this species in clinical specimens. The SSR markers developed here will enrich the limited molecular marker resource in *M. phaseolina* and could be used as the basis for more in-depth studies of the host-pathogen interactions of *M. phaseolina*.

Keywords: charcoal rot, copper sensitivity, field crops, genetic markers, *in vitro* mycelial growth, molecular fingerprinting, simple-sequence-repeat

Introduction

Macrophomina phaseolina (=*M. phaseoli*) is the causal agent of charcoal rot disease on a large number of plant species. *Macrophomina phaseolina* has been reported on four continents with a wide host range affecting more than 500 cultivated and wild plant species (Mihail & Taylor, 1995) that include economically important crops such as soybean, cotton, maize, sorghum (Cloud & Rupe, 1991; Su *et al.*, 2001), cassava (Msikita *et al.*, 1998) and sunflower (Khan, 2007). In the United States, estimated soybean yield loss to charcoal rot can reach millions of metric tons (Wrather & Koenning, 2006), whilst in other crops, such as sunflower, losses caused by charcoal rot can reach 60–90% if the conditions are favourable for infection (Khan, 2007).

Macrophomina phaseolina is an anamorphic ascomycete in the family Botryosphaeriaceae, and the absence of a known teleomorph has hindered its taxonomy over the years (Kulkarni & Patil, 1966; Crous *et al.*, 2006). In fact, it took a thorough phylogenetic study performed on 113 members of the family Botryosphaeriaceae using ribosomal DNA sequences, to be able to separate the genera *Macrophomina* and *Tiarosporella* (Crous *et al.*, 2006). Although mostly known as a plant pathogen, *M. phaseolina* is also an opportunistic human pathogen (Tan *et al.*, 2008; Srinivasan *et al.*, 2009), and so far, strains that invade plants or humans are indistinguishable from each other (Srinivasan *et al.*, 2009). As *M. phaseolina* is only present in asexual form, its identification in clinical human samples requires molecular techniques, such as DNA sequencing of the internal transcribed spacer (ITS) region (Bagyalakshmi *et al.*, 2008). This approach, however, is not ideal in clinical cases when the fast identification of this fungus is critical for timely diagnosis and treatment of infection (Srinivasan *et al.*, 2009).

Plant pathogenic isolates of *M. phaseolina* have shown variation in morphological and physiological characteristics (Pearson *et al.*, 1986a,b) as well as in the pathogenicity or host specificity (Diourte *et al.*, 1995; Mayék-Pérez *et al.*, 2001; Su *et al.*, 2001; Reyes-Franco *et al.*, 2006), thoroughly reviewed by Ndiaye (2007). Genetic variation has also been detected among *M. phaseolina* isolates using random amplified polymorphic DNA (RAPD) which identified an association between *M. phaseolina* and the plant-host origin of the isolates (Su *et al.*, 2001;

*E-mail: renee.arias@ars.usda.gov

Purkayastha *et al.*, 2006; Rajkumar & Kuruvinashetti, 2007). Similar associations were revealed by the use of universal rice primers (URP) (Jana *et al.*, 2005b) and by random amplification of fragments using oligo repeats (Jana *et al.*, 2005a; Purkayastha *et al.*, 2008; Saleh *et al.*, 2010). However, techniques such as RAPDs and URP result in randomly amplified fragments which have limited use in discovering the genetic basis of the host-pathogen associations as no genes or gene functions are determined by these methods.

Microsatellites on the other hand, also known as simple sequence repeats (SSRs), have been the most widely applied class of molecular markers used in genetic studies with applications in many fields of genetics including genetic conservation, population genetics, molecular breeding and paternity testing (Ellegren, 2004). This range of applications is due to the fact that microsatellite markers are co-dominant, multi-allelic, highly reproducible, have high-resolution, and are based on the polymerase chain reaction (PCR) (Oliveira *et al.*, 2006). As a convention, SSRs are regions in the genome where a group of bases (1–8 bp long) are repeated in tandem (Richard *et al.*, 2008). These regions can be isolated either by data mining of existing genomic sequences or by generating SSR-enriched libraries (Kijas *et al.*, 1994; Zane *et al.*, 2002).

Currently, genetic information on *M. phaseolina* is scarce with only 145 expressed sequence tags (ESTs) and 222 nucleotide sequences (46 of them containing repeats) in the National Center for Biotechnology Information (NCBI); therefore, data mining to isolate microsatellites is not an option. So far, only 13 sequence-specific microsatellites have been reported for *M. phaseolina*; these were tested on fungal isolates from soybean (Baird *et al.*, 2009) and from several other plant hosts (Baird *et al.*, 2010). In those reports, no relationships were established between the DNA sequences of the microsatellites and putative genes that could be affected by the presence of polymorphisms.

Recently, a mechanistic approach to find genetic variants associated with ecologically relevant phenotypic variation has been proposed. This approach involves predicting the impact of variations based on previous knowledge of gene function and then testing these hypotheses (Dalziel *et al.*, 2009). Thus, the use of multiple-loci sequence-specific microsatellite markers in connection with biological functions in *M. phaseolina* could provide insight towards understanding the host-pathogen relationship. In the present study a large number of microsatellites from *M. phaseolina* were isolated, the markers were tested on isolates obtained from seven different host species (Table 1), and then comparative genomics was used to gain insight into the genetic variation and possible biological processes that may contribute to specific host associations of this fungal pathogen. The results of this research show the potential of relating microsatellites to biological functions in *M. phaseolina* to elaborate testable hypotheses that can help in understanding the ecological fitness of this pathogen.

Table 1 Origin of *Macrophomina phaseolina* isolates from the USA including their corresponding hosts, location, and accession numbers at CBS-KNAW Fungal Biodiversity Center, the Netherlands

Isolate	Location	Host	CBS number[a]
TN4 (I-4280)	Jackson, TN	Soybean (*Glycine max*)	126630
TN5	Ames, TN	Soybean (*Glycine max*)	126631
TN146	Neosho, MO	Soybean (*Glycine max*)	126632
TN261	Milan, TN	Soybean (*Glycine max*)	126633
TN272	Stoneville, MS	Soybean (*Glycine max*)	126634
TN280A[b]	Jackson, TN	Soybean (*Glycine max*)	126635
TN280B	Jackson, TN	Soybean (*Glycine max*)	
TN291A	Jackson, TN	Snapbean (*Phaseolus vulgaris*)	126636
TN291B	Jackson, TN	Snapbean (*Phaseolus vulgaris*)	
TN292	Milan, TN	Sunflower (*Helianthus annuus*)	126637
TN293A	Jackson, TN	Pumpkin (*Cucurbita pepo*)	126638
TN293B	Jackson, TN	Pumpkin (*Cucurbita pepo*)	
TN294	Milan, TN	Maize (*Zea mays*)	126639
TN295A	Dyer Co., TN	Sorghum (*Sorghum bicolor*)	126640
TN295B	Dyer Co., TN	Sorghum (*Sorghum bicolor*)	
TN296	Jackson, TN	Cotton (*Gossypium hirsutum*)	126641
TN305	Madison, WI	Soybean (*Glycine max*)	126642
TN314	Cherokee Co., KS	Soybean (*Glycine max*)	126643
TN377	Dunklin Co., MO	Soybean (*Glycine max*)	126644
TN378	Stoddard Co., MO	Soybean (*Glycine max*)	126645
TN379	New Madrid Co., MO	Soybean (*Glycine max*)	126646
TN410	Jackson, TN	Sunflower (*Helianthus annuus*)	126647
TN411	Milan, TN	Sunflower (*Helianthus annuus*)	126648
TN413	Jackson, TN	Maize (*Zea mays*)	126649

[a]CBS: accession numbers assigned by CBS-KNAW Fungal Biodiversity Centre, the Netherlands. Only typical charcoal rot isolates were submitted to the CBS collection.
[b]A: typical charcoal rot characteristics, B: charcoal rot characteristics but developed a very light coloration in liquid culture. TN numbers without A or B identifiers showed typical charcoal-rot characteristics.

Materials and methods

Fungal isolates

Macrophomina phaseolina isolates were collected from each specific plant host after the crop had matured. Plant hosts included: sunflower (*Helianthus annuus*), cotton (*Gossypium hirsutum*), snap bean (*Phaseolus vulgaris*), maize (*Zea mays*), soybean (*Glycine max*), pumpkin

(*Cucurbita pepo*) and sorghum (*Sorghum bicolor*) (Table 1). Lower stems and root sections from about 10 plants were cut and thoroughly washed and dried. Dry samples were cut into 0·25 cm pieces and were placed in acidified potato dextrose agar (A-PDA) (Difco Laboratories) and incubated at 30°C for 4 days. Hyphal tip cultures were then transferred to new A-PDA plates and incubated for 4 days. The hyphal tip cultures were then transferred to a flask with 25 mL potato dextrose broth and grown on a shaker at 75 r.p.m. for 5 days. The culture suspension was filtered using vacuum suction and placed into 15 mL tubes for DNA extraction. Accession numbers assigned by CBS-KNAW are listed in Table 1.

Isolation of microsatellites

DNA for *Macrophomina phaseolina* strain TN4 (I-4280) was extracted with DNeasy Plant Maxi kit (Qiagen). SSR-enriched libraries were generated by subtraction hybridization using multiple biotinylated oligo repeats, following the protocol of Techen *et al.* (2010) briefly described here. DNA from *M. phaseolina* was digested with restriction enzymes *Alu*I, *Hae*III, *Dra*I and *Rsa*I (New England Biolabs) individually and with combinations of *Rsa*I + *Hae*III and *Dra*I + *Xmn*I. The blunt-end DNA fragments were A-tailed with Taq-DNA Polymerase (Promega) in the presence of dATP for 2 h, then ligated for 3 h at 16°C to a linker made from oligos SSRLIBF3: 5′-CGGGAGAGCAAGGAAGGAGT-3′ and SSRLIBR3: 5′Phos-CTCCTTCCTTGCTCTCTCCCGA AAA-3′ (Techen *et al.*, 2010). The ligated fragments were purified with MinElute (Qiagen) and amplified by 20 cycles of PCR using primer SSRLIBF3 and High Fidelity DNA Polymerase (Invitrogen) at: 94°C for 30 s, 60°C for 30 s and 68°C for 90 s. The amplified products were hybridized to two groups of biotinylated oligo repeats, similar to the groups listed by Glenn & Schable (2005): group 2 [$(AG)_{12}$, $(AAC)_6$, $(AAG)_8$, $(ACT)_{12}$, $(ATC)_8$], group 3 [$(AAAC)_6$, $(AAAG)_6$, $(AATC)_6$, $(AATG)_6$, $(ACAG)_6$, $(ACCT)_6$, $(ACTC)_6$, $(ACTG)_6$] and group 4 [$(AAAT)_8$, $(AACT)_8$, $(AAGT)_8$, $(ACAT)_8$, $(AGAT)_8$]; primers were purchased from Invitrogen. The final concentration of each oligo in the mix was 1 μM, and 2 μL of each oligo mix was used in 50-μL hybridization reactions. Hybridizations were performed in a gradient thermocycler at 95°C for 10 min, followed by 3 h at the annealing temperature using a gradient block at 50°C for groups 2 and 4 and 53°C for group 3, followed by an extension step of 10 min at 68°C in the presence of High Fidelity Taq Polymerase (Invitrogen) as indicated in Hayden *et al.* (2002). Sequences containing repeats were captured using streptavidin-coated magnetic beads M-270 (Invitrogen) in a Labquake tube shaker/rotator (Barnstead/Thermoline) at 22°C for 1 h (Kijas *et al.*, 1994). After binding, the beads were washed with 0·5 × SSC at ambient temperature and 0·5 × SSC at 50°C for 5 min each. Elution of the DNA from the biotinylated oligos was done with 60 μL MilliQ water at 96°C for 5 min, twice. The eluate was PCR amplified for 20 cycles as indicated for the ligation step, the PCR products were cloned in vector TOPO4 (Invitrogen) and sequenced using an ABI 3730XL DNA Analyzer (Applied Biosystems). Sequences were assembled in contigs using DNASTAR Lasergene7 (DNASTAR, Inc.) and checked visually. Repeats were searched using SSRFinder (Sharopova *et al.*, 2002) and Sputnik (available from Chris Abajian, University of Washington, WA, USA; chrisa@espresso-software.com). Primers were designed using PRIMER3 (Rozen & Skaletsky, 2000) with stringent parameter conditions: Tm 63°C optimum (min/max 60/65°C), length 24 bp optimum (min/max 20/28 bp), 3′ GC clamp, and maximum overlap of repeat within the primer was 5 bp.

Fingerprinting

Using stringent conditions in PRIMER3 software, 182 primers were designed on the flanking regions of the repeats and 24 strains of *M. phaseolina* collected from different hosts were tested (Table 1). Forward SSR primers were 5′ tailed with the sequence 5′-CAGTTTTCCCAGTCAC GAC-3′ to permit product labelling, and reverse primers were tailed at the 5′ end with the sequence 5′-GTTT-3′ to promote non-template adenylation (Brownstein *et al.*, 1996). Primer 5′-CAGTTTTCCCAGTCACGAC-3′ labelled with 6-carboxy-fluorescein (FAM) (IDT-Technologies) was used for amplification of 10 ng DNA using Titanium Taq DNA Polymerase (Clontech) in 5 μL reactions on an M&J thermal cycler (BioRad) at 95°C for 1 min, 60°C for 1 min (2 cycles), 95°C for 30 s, 60°C for 30 s, 68°C for 30 s (27 cycles) and a final extension at 68°C for 4 min. Fluorescently-labelled PCR fragments were analyzed on an ABI 3730XL DNA Analyzer and data processed using GENEMAPPER v. 3.7 (both from Applied Biosystems). Presence of alleles was converted to a binary matrix. Cluster analysis of SSR marker results for *M. phaseolina* isolates was performed using the unweighted paired group method and arithmetic averages (UPGMA) algorithm implemented in the SAHN program of NTSYSpc v. 2.2 (Exeter Software). The confidence levels for the dendrograms were assessed by bootstrap re-sampling (5000 replicates) (Felsenstein, 1985; Efron *et al.*, 1996) using WINBOOT (Yap & Nelson, 1996).

Percentage of multiallelic loci, unique pattern informative combinations (UPIC) and BLAST of contig sequences

The percentage of multiallelic loci was calculated for each accession across all the SSR markers tested. Unique pattern informative combinations (UPIC) were also calculated to identify the most informative markers and marker combinations that could discriminate all isolates evaluated. The percentage of multiallelic loci and UPIC-values were both calculated using UPIC Perl scripts (Arias *et al.*, 2009). Contig sequences containing microsatellites were screened against the NCBI Protein and Nucleotide Databases (BLASTX, BLASTNR) (Altschul *et al.*, 1990).

An additional BLAST to gene ontology was performed for all the sequences (BLAST2GO, http://www.blast2go.org).

Growth rate and copper resistance

The markers that best discriminated *M. phaseolina* isolates in the cluster analysis according to plant host were examined, and the BLASTX was analyzed for the sequences containing the discriminating microsatellites. Based on the allele distribution on those markers and BLASTX results it was decided to evaluate the growth rate and possible copper resistance in relation to host association. A subset of 16 *M. phaseolina* isolates was used for these tests (for isolates with the same name, those with typical charcoal rot characteristics were used; suffix 'A', Table 1). For growth rate, discs (6 mm diameter) of fungal colonies were placed on PDA in the centre of Petri dishes with five plates per isolate and incubated at 30°C. Radii of the colonies were measured every 12 h for a 48 h period. Data were analyzed as repeated measures using PROC MIXED and means compared by Tukey's test using SAS (Statistical Analysis System Institute).

Copper resistance of *M. phaseolina* isolates were tested by adding copper sulphate ($CuSO_4.H_2O$, Mallinckrodt) as crystals to autoclaved PDA medium to make final concentrations of 0 (control), 0·3, 0·6, 1·2, 2·4, 3·2 and 4·8 mM, and then added to Petri dishes. Five plates per copper sulphate concentration were inoculated with each isolate (1 plug of 5 mm diameter) in the centre of the dish and incubated at 30°C. When mycelial growth on control agar reached the edge of the dish, the radius of the fungal colonies was measured for all concentrations. Results were analyzed by ANOVA and means compared by Tukey's test using SAS.

Results

Fungal isolates, repeat symmetry and frequency

Macrophomina phaseolina isolates used in the present work have been stored in CBS-KNAW Fungal Biodiversity Center, the Netherlands with accession numbers CBS 126630–126649 (Table 1). SSR-enriched genomic libraries were generated from *M. phaseolina*, 864 clones were sequenced from these libraries and 531 repeats detected, 528 of which were microsatellites (1–8 bp motifs) and three were minisatellites (≥9 bp motifs). To simplify the recording of the repeat motifs, circular permutations and reverse complements of the motif sequences were grouped together as one type, i.e. AAC, ACA, CAA, GTT, TGT, TTG were recorded as AAC. The most abundant repeat motif in these libraries was AG, followed by ATC, AC and AAG. A summary of the abundance of each repeat motif is shown in Table S1. The three minisatellites detected were two 14 bp motifs (CTCACTCACACACA and ACACACACTCACTC) and one 13 bp motif (CACCACCTTCCAT), though only the latter was tested as a potential marker. A total of 680 *M. phaseolina* sequences have been submitted to GenBank, NCBI, with accession numbers GU943792-GU944472; the marker ID numbers indicated in this manuscript correspond to the contig numbers for the sequences submitted to the NCBI Database.

Marker description, multiallelic loci and UPIC values

One hundred and eighty-two primer sets were designed and tested, from which 147 amplified >90% of the 24 *M. phaseolina* isolates tested, 23 amplified <90% of the samples and 12 did not result in amplification. The group of 147 primer sets was used for cluster analysis. Primer sequences are listed in Table S2, and brief results for BLASTX (Altschul *et al.*, 1990) of their corresponding contig sequences are listed in Table S3. These 147 markers amplified between 2 and 15 alleles. The number of amplicons for each isolate was in general one; however, 50 of the markers resulted in 2–6 amplicons on individual DNA samples. Interestingly, 11 of those 50 markers had scores ≥50 on BLASTX. Overall, the 147 markers amplified 624 alleles, with an average of five alleles per marker. One interesting finding was that in seven markers (StvMPh_558skb, StvMPh_380ska, StvMPh_473, StvMPh_46, StvMPh_52, StvMPh_197 and StvMPh_45) there were isolates harbouring either allele A or allele B, while other isolates had both alleles, which may reflect the presence of heterokaryons. In general, the calculated percentage of multiallelic loci in *M. phaseolina* isolates from soybean (average 7·4%) was lower than for isolates from maize, pumpkin, snap bean and sunflower, which ranged from an average of 10·8–12·1% (Fig. 1). The total range for percentage of multiallelic loci was from a minimum of 5·6% for soybean isolate TN272 to a maximum of 16·9% for sunflower isolate TN410 (Fig. 1), indicating a variable degree of heterozygosity in the isolates.

When a large number of microsatellite markers are developed, it is useful to identify subsets of markers that can provide the maximum discrimination of DNA samples, thus reducing the cost of future experiments. The software called UPIC (unique pattern informative combination) determines the number of samples discriminated by each microsatellite marker (UPIC score) and identifies the combination of the minimum number of markers that can discriminate all the samples tested (Arias *et al.*, 2009). The UPIC software was used to identify the best markers to use in future experiments with *M. phaseolina*, and the UPIC scores different from zero were reported, as zero indicates that no sample is being uniquely discriminated by the marker (Table 2).

Cluster analysis and genetic variation of *M. phaseolina* isolates in relation to host specificity

Cluster analysis was performed for the 24 *M. phaseolina* isolates collected from seven species of plant hosts and the 147 SSR markers described above. A dendrogram summarizing the relationships among isolates is shown in Figure 1. Also listed in the figure are crop names that correspond to the crop from which the pathogens were

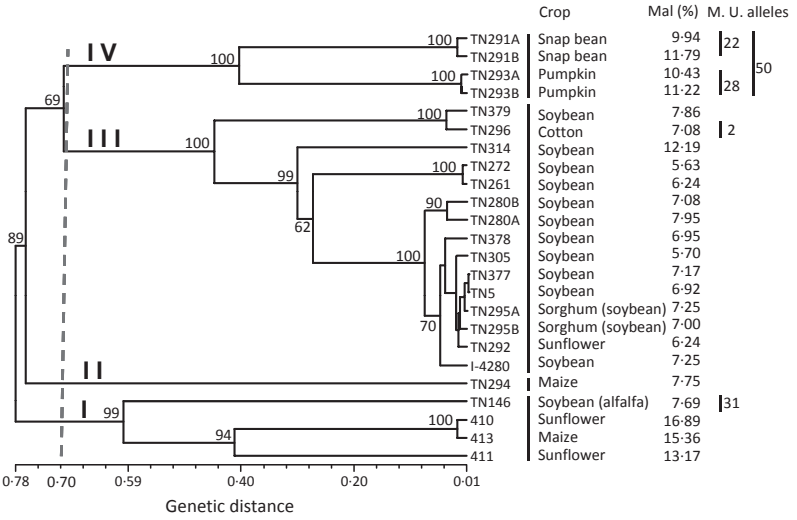

Figure 1 Cluster analysis of 24 isolates of *Macrophomina phaseolina* using 147 microsatellite markers and the unweighted paired group method and arithmetic averages (UPGMA) algorithm implemented in the SAHN program of NTSYSpc v. 2.2. Bootstrap values of 5000 resampling are shown at the nodes. MaL (%): multiallelic loci in percent, represents the percentage of loci that showed multiple amplicons. M. U. Alleles: number of markers with unique alleles found either on a particular sample or host. Dashed vertical line at an arbitrary genetic distance of 0·70 separates the four clades. In parenthesis, crops grown for several years in the plots until the season before sampling.

Table 2 Unique pattern informative combination (UPIC) values different from zero, calculated for *Macrophomina phaseolina* microsatellite markers using the UPIC software (Arias *et al.*, 2009). The eight markers indicated in bold can be used together to discriminate all 24 isolates

Marker	UPIC score	Marker	UPIC score	Marker	UPIC score	Marker	UPIC score
StvMPh_641a	5	StvMPh_52a	3	StvMPh_372a	2	StvMPh_207a	1
StvMPh_29c	4	StvMPh_49a	3	StvMPh_5a	2	**StvMPh_102a**	1
StvMPh_248a	4	StvMPh_209a	2	StvMPh_310a	2	StvMPh_461a	1
StvMPh_155a	4	StvMPh_401a	2	StvMPh_1a	2	StvMPh_378a	1
StvMPh_45a	4	StvMPh_94a	2	StvMPh_215b	2	StvMPh_289a	1
StvMPh_484a	4	StvMPh_197a	2	StvMPh_532a	1	StvMPh_16a	1
StvMPh_329a	4	StvMPh_354a	2	StvMPh_81a	1	StvMPh_226ska	1
StvMPh_173a	3	StvMPh_182a	2	StvMPh_47a	1	StvMPh_61a	1
StvMPh_247a	3	StvMPh_46a	2	StvMPh_165a	1	StvMPh_397b	1
StvMPh_507a	3	StvMPh_137a	2	StvMPh_190a	1	StvMPh_34a	1
StvMPh_25a	3	StvMPh_31b	2	StvMPh_15a	1	StvMPh_415b	1
StvMPh_232a	3	StvMPh_426a	2	**StvMPh_558a**	1	StvMPh_449a	1
StvMPh_19b	3	StvMPh_63a	2	StvMPh_352a	1	StvMPh_562c	1
StvMPh_254b	3	StvMPh_511a	2	StvMPh_578a	1	StvMPh_65a	1
StvMPh_368a	3	StvMPh_27a	2	StvMPh_280a	1	StvMPh_123a	1
StvMPh_273a	3	StvMPh_534b	2	StvMPh_146a	1	StvMPh_314a	1
StvMPh_99a	3	StvMPh_253a	2	StvMPh_380ska	1	StvMPh_430a	1
StvMPh_328a	3	StvMPh_22a	2	StvMPh_143a	1	StvMPh_473a	1
StvMPh_230a	3	**StvMPh_54a**	2	StvMPh_109a	1	StvMPh_443a	1
StvMPh_114a	3	StvMPh_144a	2	StvMPh_131a	1	StvMPh_14a	1
StvMPh_460b	3	StvMPh_92a	2	StvMPh_8a	1	StvMPh_558ska	1
StvMPh_174a	3	StvMPh_132a	2	StvMPh_558b	1	StvMPh_548a	1

isolated, the percentage of multiallelic loci and the number of markers with unique alleles shown on isolates from specific plant hosts. Three of the samples (TN146, TN295A and TN295B) were collected from crops grown for the first time in those field plots. Thus, next to the crop from where these three isolates were obtained, the crops grown for several years in those field plots until the season before sampling are indicated in parentheses. Setting an arbitrary genetic-distance threshold of 0·70 in the dendrogram identified four clusters (Fig. 1). Based on the hosts from which the fungus was isolated, Cluster I corresponded mostly to isolates from sunflower + maize, Cluster II only an isolate from maize, Cluster III was mostly isolates from soybean, and Cluster IV

corresponded to isolates from pumpkin + snap bean. High bootstrap values, 90–100%, were calculated for most nodes of the dendrogram. The clear separation was a result of the large number of markers with unique alleles for particular hosts. In some of the samples, the markers showed alleles not present in the rest of the isolates, i.e. 22 markers had alleles only present in snap bean isolates and 28 markers had alleles only present in pumpkin isolates. For isolate TN146 from 'soybean (alfalfa)' there were 31 markers with alleles not shared by any other isolate, whereas two markers had distinct alleles for the cotton sample. In addition, 50 markers showed amplicons that could discriminate pumpkin and snap bean from the rest of the isolates. No single marker could distinguish the samples labelled as 'sorghum (soybean)' from the rest of the soybean isolates. *Macrophomina phaseolina* isolates from soybean all grouped into Cluster III with the exception of TN146. Markers that showed distinct amplicon patterns among the four clusters are shown in Table 3.

Forty SSR markers described here had BLASTX-score values of 50 or higher when their contig sequences were compared to NCBI databases (Table S3). At the same time, many of these SSR markers discriminated isolates by plant host. Biological processes deduced from the marker sequences using BLAST2GO (2009) included morphogenesis, solute transport, apoptosis, signal transduction, oxidation reduction and biosynthesis (data not shown). Essentially, the genetic variation observed could be related to fungal physiology and host origin of the *M. phaseolina* isolates (Table S3).

Growth rate and copper resistance

Some of the SSR markers that best distinguished the four main clusters in the cluster analysis were related to growth rate and copper resistance (Fig. 1, Table 3) and these responses were examined in more detail. For example, samples from Cluster IV showed the slowest growth rate (Table 4a), and had unique alleles 120 and 171 for markers StvMPh_146a and StvMPh_310a, respectively (Table 3). Samples from Clusters II and IV had the lowest resistance to copper (Table 4b) and had alleles 107 and 112 for marker StvMPh_182a, not present on other clusters (Table 3). Among these markers, StvMPh_146a and StvMPh_310a corresponded to Snd1/p100-transcription factor and a nuclear elonga-

Table 3 Markers that most clearly distinguished the four clusters of isolates of *Macrophomina phaseolina* indicated in the dendrogram. Numbers in each cluster column are the allele sizes for the corresponding locus

	Alleles present						
Locus	Cluster I	Cluster II	Cluster III	Cluster IV	BLASTX and BLAST2GO	Score	Expected
StvMPh_310a	165	168	169, 175	171	X: Nuclear elongation and deformation protein, lipin Smp2, gb|EEH09799·1| 2GO: no GO terms	139	8·00E-32
StvMPh_461a	117, 117 + 173, 173 + 122	122	115	95	X: Pisatin demethylase, cytochrome p450, ref|XP_001259930·1| 2GO: Oxidation reduction; metal ion binding, methyltransferase activity	138	1·08E-51
StvMPh_146a	100	100	112	120	X: Snd1/p100 transcription factor, gb|EDP51668·1| 2GO: RNA interference; nucleic acid binding, hydrolase activity	91	1·67E-15
StvMph_562c	121, 145 179	182	177	168	X: ref|XP_001263010·1| enoyl-CoA hydratase/isomerase family protein 2GO: Metabolic process; catalytic activity, lyase activity, isomerase	78	4·90E-06
StvMPh_63a	165	157	161, 171, 177	165, 173	X: NmrA-like family protein, gb|EED12072·1| 2GO: Metabolic process; catalytic activity, binding	49·3	1·00E-04
StvMPh_415b	186, 202, 235	170	182	181, 184	X: low score (riboflavin biosynthesis) 2GO: no GO terms		
StvMPh_182a	103, 108, 135	107	106, 116, 118	112	X: low score (copper oxidase) 2GO: no GO terms		

Table 4 Mean colony radius (n = 5) of *Macrophomina phaseolina* isolates on (a) potato dextrose agar after 48 h incubation and (b) 2·4 and 3·2 mM $CuSO_4.H_2O$

(a) Cluster[a]	Isolate	Colony radius (cm)	Groups[b] $\alpha \leq 0{\cdot}01$	(b) Cluster	Isolate	Copper resistance radius (cm)	Groups $\alpha \leq 0{\cdot}01$
III	TN261	4·04	a	III	TN261	4·02	a
I	TN146	3·58	ab	III	TN292	3·88	ab
III	TN272	3·50	b	III	TN272	3·58	bc
III	TN280	3·50	b	III	TN305	3·56	bc
III	TN378	3·50	b	I	TN146	3·45	bc
III	TN292	3·48	b	III	TN379	3·35	cd
III	TN377	3·46	b	III	TN280	3·16	cde
III	TN314	3·46	b	III	TN377	3·01	def
III	TN379	3·42	b	III	TN5	3·00	def
II	TN294	3·40	b	III	TN314	2·95	def
III	TN296	3·32	b	III	TN296	2·91	ef
III	TN295	3·28	bc	III	TN295	2·91	ef
III	TN305	3·28	bc	III	TN378	2·79	efg
III	TN5	3·16	bc	IV	TN291	2·69	fg
IV	TN291	2·78	cd	IV	TN293	2·44	gh
IV	TN293	2·48	d	II	TN294	2·24	h

[a]Cluster I, II, III IV: indicates the cluster where isolates were assigned in the dendrogram (Fig. 1).
[b]Groups: same letters indicate no statistically significant differences at $\alpha \leq 0{\cdot}01$.

tion protein, both related to cell growth (Tong *et al.*, 1995; Shaw *et al.*, 1998). To further investigate a possible association between these two markers and growth rate, a subset of 16 isolates were evaluated for their growth on PDA medium at 12 h interval samplings using five technical replicates. The overall colony radius of the isolates after 48 h incubation is shown in Table 4. A gradient of growth rates was observed for the isolates, with the fastest being TN261 and TN146, isolated from soybean, and the slowest TN293 isolated from pumpkin ($P \leq 0{\cdot}01$), Table 4. On average, isolates from Clusters I and III had faster growth rates than the isolates in Cluster IV (Table 4). It remains to be tested if alleles present in Cluster III isolates for markers StvMPh_146a and StvMPh_310a are also present in other fast growing isolates.

Another SSR marker that clearly distinguished the various clusters was StvMPh_182a, which had a low score homology to a copper oxidase or copper resistance protein. Copper resistance was analysed in a subset of 16 *M. phaseolina* isolates to find possible relationships between the marker and the physiology of the pathogen. These isolates were grown on seven copper sulphate concentrations (0, 0·3, 0·6, 1·2, 2·4, 3·2 and 4·8 mM). In a previous study with the dimorphic fungus *Mucor rouxii*, copper sulphate concentrations of 3·2, 4·8 and 6·0 mM were considered to be high, and strains whose spores could germinate and grow on 6·0 mM were considered resistant (Ramirez-Salgado *et al.*, 1996). Growth at copper sulphate concentrations of 0·3 and 0·6 mM did not significantly differ from the copper-free control on any of the *M. phaseolina* isolates ($\alpha \leq 0{\cdot}05$). None of the isolates grew at 4·8 mM copper sulphate; results for concentrations of 2·4 and 3·2 mM combined are shown in Table 4. The most resistant isolates were TN261 and TN292 ($P \leq 0{\cdot}01$) located in Cluster III and the most sensitive isolates were TN291, TN293 and TN294 ($P \leq 0{\cdot}01$) located in Clusters II and IV (Table 4). In addition to StvMPh_182a, another two markers that would be interesting to analyze in relation to copper resistance, both by biological process and allele sizes were StvMph_94a and StvMph_116a. These markers had similarity to a cation-diffusion facilitator and a TRAP-transporter solute receptor, respectively (Table S3). Cluster III isolates had alleles 107 + 206 or 104 + 202 for marker StvMPh_94a, which were not present in the other clusters; also alleles 134 or 137 for marker StvMPh_116a were not present in other clusters except for isolate TN146. Though no direct association of alleles and copper resistance can be made at this time, in general isolates in Clusters I and III had more resistance to copper than Clusters II and IV, and these clusters were distinguished by three markers related to copper oxidation or cation transport. The overall trend in the level of resistance to copper sulphate of the isolates is shown in Table 4.

Discussion

Single alleles per marker on each sample have been reported previously when 13 microsatellite markers were tested on *M. phaseolina* isolates from soybean (Baird *et al.*, 2009, 2010) and from other plant hosts (Baird *et al.*, 2010). In the present work, however, 30% of the microsatellite markers analyzed revealed more than one allele on at least one of the *M. phaseolina* isolates obtained from various plant hosts, including 12 isolates collected from soybean. Interestingly, 22% of the markers that were multiallelic on individual samples also had high score hits on BLASTX, indicating potential genetic variability on actual proteins. The percentage of multiallelic loci detected by SSRs in an organism is usually correlated to the overall heterozygosity of its genome (Aparicio *et al.*, 2007); in addition, high heterozygosity is considered as high potential for adaptation to the environment (Hansson & Westerberg, 2002). In general, isolates from soybean had a lower percentage of multiallelic loci than the rest of the isolates, with the exception of TN314 (Fig. 1). Although future experiments will be needed to analyze the nuclear state of these isolates, if this trend remains, this could indicate a low level of genetic exchange among soybean isolates, perhaps because this host association is probably dominated by a more homogeneous population of *M. phaseolina*. Further work is required to test for the presence/impact of heterokaryons and diploids and their potential role on the ecological fitness of this pathogen.

Currently, the genetic resources for *M. phaseolina* in public databases are limited. The present work has contributed 680 nucleotide sequences of *M. phaseolina*, 220 of these corresponding to microsatellites that were

mostly polymorphic across the isolates tested. As previously mentioned, accurate identification of *M. phaseolina* can be difficult and may require molecular systematics (Crous *et al.*, 2006). In a case study of a human infection with *M. phaseolina*, cultural methods were time consuming and not sufficient for identification, thus requiring analysis of ITS and 28S rDNA sequences (Srinivasan *et al.*, 2009). The present work identified 27 monomorphic markers that amplified across all isolates evaluated. These markers are indicated in bold in Table S2. If these markers prove to be universal for *M. phaseolina* they could probably be multiplexed to develop a PCR assay for the fast identification of this pathogen in human infections.

The unique pattern informative combination (UPIC) score of a marker corresponds to the number of DNA samples that can be individually discriminated from the rest of the samples by that particular marker, i.e. a UPIC value of 2 indicates that two samples have unique allele patterns for that marker and can be distinguished from each other and from the rest of the samples (Arias *et al.*, 2009). The present work lists 80 markers with UPIC scores different from zero. These can be used to plan future experiments, allowing the selection of the most informative SSR markers developed, thus reducing the cost of the analyses.

For many years, morphological, physiological and genetic variations have been observed in *M. phaseolina*. In some cases those variations have even been associated with the plant host origin of the isolates (Pearson *et al.*, 1986a,b; Diourte *et al.*, 1995; Su *et al.*, 2001; Purkayastha *et al.*, 2006, 2008; Rajkumar & Kuruvinashetti, 2007; Baird *et al.*, 2010); however, the methods used did not allow identification of potential genes that could be involved in the pathogen-host association. Performing BLASTX and gene ontology of the sequences showed that some aspects of the physiology of *M. phaseolina* could have been altered in relation to the association with the host from which each isolate was collected. *Macrophomina phaseolina* isolates from soybean all grouped into Cluster III despite having been collected from five different states (MS, MO, TN, KS and WI). The only exception was TN146, the 'soybean (alfalfa)' isolate that grouped with Cluster I. In general, Cluster I grouped isolates from sunflower and maize, and cluster IV corresponded to isolates from pumpkin and snap bean.

From the group of markers that most clearly separated the four clusters in the dendrogram (Fig. 1, Table 3), the three markers with the highest BLASTX scores corresponded to a pisatin demethylase-cytochrome P450, a transcription factor involved in RNA interference, and a nuclear elongation and deformation factor. The marker associated with pisatin demethylase-P450 had a unique allele for isolates from pumpkin and snap bean, and another unique allele for soybean isolates. It has been shown in *Nectria haematococca* that demethylation of pisatin is encoded by genes at three different loci, and highly active demethylation is linked to pathogenicity (Kistler & VanEtten, 1984). It is also known that the cytochrome P450 component determines the substrate specificity of the demethylation reaction, and that the level of expression of its associated NADPH-cytochrome-P450 reductase could affect pathogenicity (Desjardins & VanEtten, 1986), and play a major role in adaptation to ecological niches (Deng *et al.*, 2007).

A second marker with a high BLASTX score that clearly distinguished the clusters in the dendrogram matched Snd1/p100 transcription factor. Snd1 refers to *Staphylococcus* nuclease domain containing 1, which harbours multiple domains. It was first described as transcription activator (Tong *et al.*, 1995), then other functions were described for Snd1/p100, such as chromatin modification by recruiting histone acetyltransferase activity (Valineva *et al.*, 2005), and RNA interference (Yang *et al.*, 2007). Since single and unique alleles were observed for this marker within each cluster, it would be interesting to analyze the protein function in relation to the plant host origin of the isolate.

Other markers that distinguished the clusters had low BLASTX scores with some homology to riboflavin biosynthesis, regulation of nitrogen utilization, an isomerase and a copper oxidase. The marker related to the latter had three alleles that could be related to the higher resistance to copper observed in the soybean isolates (Cluster III, Fig. 1). Fungi have developed numerous mechanisms for copper resistance (Cervantes & Gutierrez-Corona, 1994) some of them involving cell wall composition, and in some fungi copper resistance has been related to greater ability of spore germination (Ramirez-Salgado *et al.*, 1996). This is another aspect that would be interesting to explore among the host associations of *M. phaseolina*.

The objectives of this work were to isolate a large number of sequence-specific microsatellites from *M. phaseolina*, screen them on isolates from various plant hosts to identify polymorphisms, and examine the potential association of the polymorphisms to biological processes in the pathogen. Thus, this study fulfilled various purposes. First, only 13 markers had previously been reported for *M. phaseolina* (Baird *et al.*, 2009), and while studies using few loci can infer recent population history, the analysis of large number of loci is more likely to contain information about older events (Wang & Hey, 2010). The present work provides 147 new microsatellite markers that will significantly enrich the molecular resources for *M. phaseolina*. Secondly, by calculating UPIC scores, markers were identified that detected the maximum genetic variations on *M. phaseolina* samples, hence minimizing cost of future population studies. Thirdly, Dalziel *et al.* (2009) had recommended a mechanistic analysis of molecular markers in relation to gene function as a more effective way to understand environmental fitness. Performance of BLASTX and BLAST2GO on the sequences allowed examination for the first time of biological processes potentially affected by polymorphisms at the loci that best distinguished the plant host origin of the isolates of *M. phaseolina*. The comprehensive approach used in the present work, and

the large number of microsatellite markers reported, will allow elaboration of testable hypotheses regarding heterokaryons, ploidy, and environmental fitness of this fungus and provide insights into the host-pathogen association in the plant-*M. phaseolina* interaction.

Acknowledgements

This work was supported by USDA-ARS project numbers 6402-21220-010-00D, 6401-21220-001-00D and 6402-21310-003-00. Additional support was provided by United Soybean Board project number 8245. We thank Ms Xiaofen (Fanny) Liu for library sequencing and Ms Sheron A. Simpson for testing the SSR markers. Any library/sequence information requirements can be addressed to Dr Brian E. Scheffler at brian.scheffler@ars.usda.gov. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ, 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–10.

Aparicio JM, Ortego J, Cordero PJ, 2007. Can a simple algebraic analysis predict markers-genome heterozygosity correlations? *Journal of Heredity* **98**, 93–6.

Arias RS, Ballard LE, Scheffler BE, 2009. UPIC: Perl scripts to determine the number of SSR markers to run. *Bioinformation* **3**, 353–61.

Bagyalakshmi R, Therese KL, Prasanna S, Madhavan HN, 2008. Newer emerging pathogens of ocular non-sporulating molds (NSM) identified by polymerase chain reaction (PCR)-based DNA sequencing technique targeting internal transcribed spacer (ITS) region. *Current Eye Research* **33**, 139–47.

Baird RE, Wadl PA, Wang X *et al.*, 2009. Microsatellites from the charcoal rot fungus *Macrophomina phaseolina*. *Molecular Ecology Resources* **9**, 946–8.

Baird RE, Wadl PA, Allen T *et al.*, 2010. Variability of United States isolates of *Macrophomina phaseolina* based on simple sequence repeats and cross genus transferability to related genera within Botryosphaeriaceae. *Mycopathologia* **170**, 169–80.

Brownstein MJ, Carpten JD, Smith JR, 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *BioTechniques* **20**, 1004–10.

Cervantes C, Gutierrez-Corona F, 1994. Copper resistance mechanisms in bacteria and fungi. *FEMS Microbiology Reviews* **14**, 121–37.

Cloud GL, Rupe JC, 1991. Morphological instability on a chlorate medium of isolates of *Macrophomina phaseolina* from soybean and sorghum. *Phytopathology* **81**, 892–5.

Crous PW, Slippers B, Wingfield MJ *et al.*, 2006. Phylogenetic lineages in the Botryosphaeriaceae. *Studies in Mycology* **55**, 235–53.

Dalziel AC, Rogers SM, Schulte PM, 2009. Linking genotypes to phenotypes and fitness: how mechanistic biology can inform molecular ecology. *Molecular Ecology* **18**, 4997–5017.

Deng J, Carbone I, Dean RA, 2007. The evolutionary history of cytochrome P450 genes in four filamentous Ascomycetes. *BMC Evolutionary Biology* **7**, 30.

Desjardins AE, VanEtten HD, 1986. Partial purification of pisatin demethylase, a cytochrome P-450 from the pathogenic fungus *Nectria haematococca*. *Archives of Microbiology* **144**, 84–90.

Diourte M, Starr JL, Jeger MJ, Stack JP, Rosenow DT, 1995. Charcoal rot (*Macrophomina phaseolina*) resistance and the effects of water-stress on disease development in sorghum. *Plant Pathology* **44**, 196–202.

Efron B, Halloran E, Holmes S, 1996. Bootstrap confidence levels for phylogenetic trees. *Proceedings of the National Academy of Sciences, USA* **93**, 13429–34.

Ellegren H, 2004. Microsatellites: simple sequences with complex evolution. *Nature Reviews. Genetics* **5**, 435–45.

Felsenstein J, 1985. Confidence-limits on phylogenies – an approach using the bootstrap. *Evolution* **39**, 783–91.

Glenn TC, Schable NA, 2005. Isolating microsatellite DNA loci. *Methods in Enzymology* **395**, 202–22.

Hansson B, Westerberg L, 2002. On the correlation between heterozygosity and fitness in natural populations. *Molecular Ecology* **11**, 2467–74.

Hayden MJ, Good G, Sharp PJ, 2002. Sequence tagged microsatellite profiling (STMP): improved isolation of DNA sequence flanking target SSRs. *Nucleic Acids Research* **30**, e129.

Jana T, Sharma TR, Singh NK, 2005a. SSR-based detection of genetic variability in the charcoal root rot pathogen *Macrophomina phaseolina*. *Mycological Research* **109**, 81–6.

Jana TK, Singh NK, Koundal KR, Sharma TR, 2005b. Genetic differentiation of charcoal rot pathogen, *Macrophomina phaseolina*, into specific groups using URP-PCR. *Canadian Journal of Microbiology* **51**, 159–64.

Khan SN, 2007. *Macrophomina phaseolina* as causal agent for charcoal rot of sunflower. *Mycopathology* **5**, 111–8.

Kijas JM, Fowler JC, Garbett CA, Thomas MR, 1994. Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *BioTechniques* **16**, 656–60.

Kistler HC, VanEtten HD, 1984. Three non-allelic genes for pisatin demethylation in the fungus *Nectria haematococca*. *Journal of General Microbiology* **130**, 2595–603.

Kulkarni NB, Patil BC, 1966. Taxonomy and discussion on the nomenclature of *Macrophomina phaseoli* (Maubl.) Ashby, and its isolates from India. *Mycopathologia* **28**, 257–64.

Mayék-Pérez N, López-Castañeda C, González-Chavira M *et al.*, 2001. Variability of Mexican isolates of *Macrophomina phaseolina* based on pathogenesis and AFLP genotype. *Physiological and Molecular Plant Pathology* **59**, 257–64.

Mihail JD, Taylor SJ, 1995. Interpreting variability among isolates of *Macrophomina phaseolina* in pathogenicity, pycnidium production, and chlorate utilization. *Canadian Journal of Botany* **73**, 1596–603.

Msikita W, James B, Wilkinson HT, Juba JH, 1998. First report of *Macrophomina phaseolina* causing pre-harvest cassava root rot in Benin and Nigeria. *Plant Disease* **82**, 1402.

Ndiaye M, 2007. *Ecology and Management of Charcoal Rot* (Macrophomina phaseolina) *on Cowpea in the Sahel*. Wageningen, the Netherlands: Wageningen University, PhD thesis.

Oliveira EJ, Pádua JG, Zucchi MI, Vencovsky R, Vieira MLC, 2006. Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology* **29**, 294–307.

Pearson CAS, Leslie JF, Schwenk FW, 1986a. Variable chlorate resistance in *Macrophomina phaseolina* from corn, soybean, and soil. *Phytopathology* **76**, 646–9.

Pearson CAS, Leslie JF, Schwenk FW, 1986b. Differential catabolism of nitrogen by chlorate resistant and sensitive isolates of *Macrophomina phaseolina*. *Phytopathology* **76**, 1145.

Purkayastha S, Kaur B, Dilbaghi N, Chaudhury A, 2006. Characterization of *Macrophomina phaseolina*, the charcoal rot pathogen of cluster bean, using conventional techniques and PCR-based molecular markers. *Plant Pathology* **55**, 106–16.

Purkayastha S, Kaur B, Arora P, Bisyer I, Dilbaghi N, Chaudhury A, 2008. Molecular genotyping of *Macrophomina phaseolina* isolates: comparison of microsatellite primed PCR and repetitive element sequence-based PCR. *Journal of Phytopathology* **156**, 372–81.

Rajkumar BF, Kuruvinashetti MS, 2007. Genetic variability of sorghum charcoal rot pathogen (*Macrophomina phaseolina*) assessed by random DNA markers. *The Plant Pathology Journal* **23**, 45–50.

Ramirez-Salgado P, Caracheo-Aguilar S, Obregon-Herrera A *et al.*, 1996. Copper resistance and accumulation in the zygomycete *Mucor rouxii*. *Current Microbiology* **33**, 163–6.

Reyes-Franco MC, Hernandez-Delgado S, Beas-Fernandez R, Medina-Fernandez M, Simpson J, Mayek-Perez N, 2006. Pathogenic and genetic variability within *Macrophomina phaseolina* from Mexico and other countries. *Journal of Phytopathology* **154**, 447–53.

Richard GF, Kerrest A, Dujon B, 2008. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiology and Molecular Biology Reviews* **72**, 686–727.

Rozen S, Skaletsky H, 2000. Bioinformatics methods and protocols. In: Misener S, Krawetz S, eds. *Methods in Molecular Biology*, Vol. **132**. Totowa, NJ, USA: Humana Press, 365–86.

Saleh AA, Ahmed HU, Todd TC *et al.*, 2010. Relatedness of *Macrophomina phaseolina* isolates from tallgrass prairie, maize, soybean and sorghum. *Molecular Ecology* **19**, 79–91.

Sharopova N, McMullen MD, Schultz L *et al.*, 2002. Development and mapping of SSR markers for maize. *Plant Molecular Biology* **48**, 463–81.

Shaw SL, Maddox P, Skibbens RV, Yeh E, Salmon ED, Bloom K, 1998. Nuclear and spindle dynamics in budding yeast. *Molecular Biology of the Cell* **9**, 1627–31.

Srinivasan A, Wickes BL, Romanelli AM *et al.*, 2009. Cutaneous infection caused by *Macrophomina phaseolina* in a child with acute myeloid leukemia. *Journal of Clinical Microbiology* **47**, 1969–72.

Su G, Suh S-O, Schneider RW, Russin JS, 2001. Host specialization in the charcoal rot fungus, *Macrophomina phaseolina*. *Phytopathology* **91**, 120–6.

Tan DHS, Sigler L, Gibas CFC, Fong IW, 2008. Disseminated fungal infection in a renal transplant recipient involving *Macrophomina phaseolina* and *Scytalidium dimidiatum*: case report and review of taxonomic changes among medically important members of the Botryosphaeriaceae. *Medical Mycology* **46**, 285–92.

Techen N, Arias RS, Glynn NC, Pan Z, Khan I, Scheffler BE, 2010. Optimized construction of microsatellite-enriched libraries. *Molecular Ecology Resources* **10**, 508–15.

Tong X, Drapkin R, Yalamanchili R, Mosialos G, Kieff E, 1995. The Epstein-Barr virus nuclear protein 2 acidic domain forms a complex with a novel cellular coactivator that can interact with TFIIE. *Molecular and Cellular Biology* **15**, 4735–44.

Valineva T, Yang J, Palovuori R, Silvennoinen O, 2005. The transcriptional co-activator protein p100 recruits histone acetyltransferase activity to STAT6 and mediates interaction between the CREB-binding protein and STAT6. *The Journal of Biological Chemistry* **280**, 14989–96.

Wang Y, Hey J, 2010. Estimating divergence parameters with small samples from a large number of loci. *Genetics* **184**, 363–79.

Wrather JA, Koenning SR, 2006. Estimates of disease effects on soybean yields in the United States 2003 to 2005. *Journal of Nematology* **38**, 173–80.

Yang J, Valineva T, Hong J *et al.*, 2007. Transcriptional co-activator protein p100 interacts with snRNP proteins and facilitates the assembly of the spliceosome. *Nucleic Acids Research* **35**, 4485–94.

Yap IV, Nelson RJ, 1996. WinBoot: a program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms. *IRRI Discussion Paper Series No. 14*. Manila, Philippines: International Rice Research Institute.

Zane L, Bargelloni L, Patarnello T, 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology* **11**, 1–16.

Supporting information

Additional supporting information may be found in the online version of this article:

Table S1. Motifs and motif frequencies detected on microsatellite-enriched libraries of *Macrophomina phaseolina*.

Table S2. Microsatellite markers from *Macrophomina phaseolina*. Marker IDs in bold font were monomorphic for the 24 isolates tested. Markers whose DNA sequences had significant scores on BLASTX and/or BLAST2GO are indicated in shaded rows.

Table S3. BLAST results for the sequences from which the microsatellite markers were isolated.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset by Wiley-Blackwell. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.