The development of a novel diagnostic PCR for *Madurella mycetomatis* using a comparative genome approach

Wilson Lim[1], Kimberly Eadie[1], Emmanuel Siddig[2], Bertrand Nyuykonge[1], Sarah Ahmed[3], Ahmed H. Fahal[2], Annelies Verbon[1], Sandra Smit[4], Wendy WJ van de Sande[1]

[1]Erasmus MC, University Medical Center Rotterdam, Department of Microbiology and Infectious Diseases, Rotterdam, The Netherlands.

[2]Mycetoma Research Centre, University of Khartoum, Khartoum, Sudan

[3]Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands

[4]Bioinformatics Group, Wageningen University & Research, Wageningen, The Netherlands

Corresponding author

Wendy WJ van de Sande

w.vandesande@erasmusmc.nl

Keyword: Mycetoma, *Madurella mycetomatis*, Comparative genome, Neglected tropical disease

ABSTRACT:

Eumycetoma is a neglected tropical disease characterized by large tumorous lesions. It is most commonly caused by the fungus *Madurella mycetomatis* which accounts for more than 70% of cases in central Africa. Currently, identification of the causative agent can only be reliably performed by a species-specific PCR. However, we recently demonstrated that our *M. mycetomatis* specific PCR can cross-react with *Madurella pseudomycetomatis.* We therefore used a comparative genome approach to develop a new *M. mycetomatis* specific PCR for species identification. For this we compared the published *M. mycetomatis* genome to genomes of other organisms in BLASTCLUST to identify unique *M. mycetomatis* predicted protein coding sequences. Based on 16 of these unique sequences, PCR primers were developed. The specificity of these primers was further evaluated in other eumycetoma causing agents including the *Madurella* sibling species. Out of the 16 tested sequences, only one was unique for *M. mycetomatis* and this should be used as a novel diagnostic marker for *M. mycetomatis.*

INTRODUCTION:

The neglected tropical disease mycetoma presents itself as a subcutaneous chronic granulomatous infectious and inflammatory disease and is characterized by tumorous lesions (1, 2). This disease can be caused by more than 70 different micro-organisms and is categorized into actinomycetoma (caused by bacteria) and eumycetoma (caused by fungi). Most cases occur in the mycetoma belt between the latitudes 15° South and 30° North. Diagnosis of eumycetoma is often only made clinically in endemic areas due to the scarcity of facilities, expertise and financial capacity. Identification of the causative agent is time consuming and often limited to culture and histology which can leads to misidentifications (1, 3, 4). The only way to properly identify eumycetoma causative agents to the species level is through molecular identification, commonly PCR. PCR identification's reliability, high turnover and sensitivity has made it widely used in diagnosing fungal disease, and in most cases, it has already replaced culturing of the microorganism as the primary diagnostic method (5-11).

Madurella mycetomatis is a fungus only recognised as a pathogen in mycetoma and is responsible for more than 70% of all mycetoma infections in endemic areas (1, 12, 13). In 1999, specific PCR primers based on the internal transcribed spacer (ITS) region were developed for *M. mycetomatis* by our group (14), however, recently, this *M. mycetomatis* specific PCR primer pair was discovered to cross-react with *Madurella pseudomycetomatis* (15). Back then, *M. pseudomycetomatis* was not yet discovered (16). Since new fungi causing eumycetoma are still being discovered, there is clearly a need for a specific PCR marker to identify *M. mycetomatis*. *M. pseudomycetomatis* are more commonly found in Central and South America, while *M. mycetomatis* is predominant in the African continent (17). All four *Madurella* species (*Madurella fahalii, Madurella tropicana, M. mycetomatis* and *M. pseudomycetomatis*) are known to cause mycetoma and requires different treatment strategies, thus, it is important to be able to distinguish between them in order to study their

epidemiology and to administer proper treatment (18). Since all four *Madurella* species share a very conserved ITS region, this has made designing PCR primers specific for *M. mycetomatis* based on that region difficult (18, 19). To avoid the complication brought by the ITS region, we took a different approach to design specific primers to diagnose *M. mycetomatis* using PCR. Here, using the unique conserved DNA sequence from the published genome of *M. mycetomatis* (20), we have identified diagnostic DNA markers and developed a new species specific PCR to diagnose *M. mycetomatis.*

MATERIALS AND METHOD:

Fungal isolates

A total of 93 fungal isolates were used in this study; 60 *M. mycetomatis*, 4 *M. tropicana,* 3 *M. fahalii,* 3 *M. pseudomycetomatis,* 1 *Aspergillus fumigatus,* 1 *Aspergillus terreus,* 2 *Chaetomium globosum,* 4 *Falciformispora senegalensis,* 1 *Fusarium solani,* 3 *Medicopsis romeroi,* 3 *Thielavia terrestris,* 3 *Thielavia subthermophilia,* 4 *Trematospheria grisea* and 1 *Trichophyton rubrum.* These fungal isolates were obtained from both the Mycetoma Research Center in Sudan and the Westerdijk Fungal Biodiversity Institute in the Netherlands and maintained in Erasmus Medical Centre. All isolates were identified to the species level on the basis of morphology, polymerase chain reaction (PCR)-based restriction fragment length polymorphisms, and sequencing of the ITS regions (3, 14, 21).

DNA isolation

Fungal isolates were first grown on Sabouraud Dextrose agar (Difco Laboratories) with or without gentamicin (40mg/mL, Centrafarm Nederland B.V., The Netherlands) for three weeks at either 37°C or room temperature depending on the fungal species. The mycelium was then harvested from the agar and subjected to 3 minutes of bead bashing (Tissuelyser, Qiagen) at maximum power of 30 frequency per second using ten to twelve 3mm metal beads (DIT Holland B.V., The Netherlands) per 2mL Eppendorf tube. DNA was then isolated using ZR Fungal/Bacterial DNA MicroPrepTM kit (Zymo

Research, Irvine, California, USA) according to manufacturer instructions while omitting the bead bashing step from the kit.

Identifying unique predicted protein coding sequences in the *M. mycetomatis* genome

M. mycetomatis predicted protein coding sequences were obtained from the recently published genome sequence of *M. mycetomatis* isolate mm55, accession number LCTW00000000, BioProject PRJNA267680 (20). To determine their specificity to *M. mycetomatis*, a bioinformatical comparison between the predicted protein coding sequences and the genome of other organisms was performed using BLASTCLUST (22). Predicted protein coding sequences were chosen depending on their orthologues, E-value and fragment size. Orthologues were defined as predicted amino acid sequences with greater than 85% amino acid identity to *M. mycetomatis* proteins. Predicted protein coding sequences with no orthologues present in the genomes of other organisms was chosen based on the E-value determined through BLASTCLUST. Predicted protein coding sequences with an E-value of 0.003 and higher was selected as a cut-off point and considered to be specific to *M. mycetomatis* and those between 400 bp and 1100 bp was further selected for analysis.

Primer design and PCR conditions:

Forward and reverse primers were designed according to the nucleotide sequence of the predicted protein coding sequences of interest. Primer sequences are depicted in Table 1. PCR reaction was set up with to a final volume of 25µl containing 0.6 units of Super Taq HC DNA polymerase (Sphaero Q), 0.1 nM/µl DNTP (Thermo Fisher Scientific) and 0.5 pmol/µl of each forward and reverse primer. DNA was amplified in a thermal cycler (Applied Biosystems VeritiTM) using the following program: initial denaturation at 94°C for 10 min; 40 cycles of amplification with various annealing temperatures (95°C for 1 minute, 55-59°C for 1 minute, and 72°C for 1 minute); and a final extension step of 10 seconds at 72°C. The PCR reaction products were visualized using 2% agarose gel (Sphaero Q) with GeneRuler 100bp Plus DNA ladder (Thermo Fisher Scientific) and stained with SYBR® Safe DNA Gel

Stain (Thermo Fisher Scientific). Bands on gel were visualized using a gel imaging machine (Isogen Life Science B. V., The Netherlands).

RESULTS and DISCUSSION

Recently, the currently used *M. mycetomatis* specific PCR primers were discovered to cross-react with *M. pseudomycetomatis* (15). To be able to study the epidemiology of the fungi and to administer proper treatment on patients, it is important to be able to distinguish between the different fungal species that cause eumycetoma, therefore, there is a need to design *M. mycetomatis* specific primers for diagnostic purposes. From the genome of *M. mycetomatis,* a total of 350 predicted protein coding sequences were selected and analyzed. We chose to select predicted protein coding sequences because protein coding regions are likely to be more stable than non-coding (23, 24). To ensure that they can be easily amplified through PCR, we preferentially chose predicted protein coding sequences with sizes between 400 and 1100bp. From the initial 350 predicted protein coding sequences, the top 16 candidates that fitted our requirement were chosen for PCR development.

PCR primers for the 16 candidates were then designed (Table 1). To ascertain that the primers would amplify their targets in all *M. mycetomatis* isolates, the PCR primers were evaluated in 60 *M. mycetomatis* isolates from different geographical origin, genotypic background and phenotypic appearance. Out of the 16 primer sets tested, 13 were positive in all *M. mycetomatis* isolates (figure 1). Primer sets 4, 5 and 12 are present in 58, 4 and 59 isolates, respectively (figure 1). To determine the specificity of the 13 positive primer sets, they were tested against other fungal mycetoma causative agents and close relatives of *M. mycetomatis*. As seen in table 2, only primer set 11 – later renamed as Mmy-Fw and Mmy-Rv - was found to be specific for *M. mycetomatis*. Primer set 2, 4, 8 and 9 were not able to discriminate between the different *Madurella* species while 5 and 7 could

discriminate between the four *Madurella* species but cross reacted with at least one other mycetoma causative agent. Mmy-Fw and Mmy-Rv appears to be a putative single-copy gene, making it an ideal candidate as an identification marker. Since it seemed the most suited for identifying *M. mycetomatis,* it was compared to the currently used diagnostic PCR to determine the limit of detection. Mmy-Fw and Mmy-Rv was observed to be only slightly less sensitive compared to the current *M. mycetomatis* specific PCR primer pair 26.1a and 28.3a. Mmy-Fw and Mmy-Rv was able to detect DNA concentrations as low as 5 pg whereas the slightly more sensitive *M. mycetomatis* specific PCR was able to detect DNA at 0.5 pg.

One of the advantages of the comparative genome method is that primer design is less constrained since the targeted genes are unique. With this method, we were able to design primers that are able to distinguish between *M. mycetomatis* and *M. pseudomycetomatis.* Using this approach, other studies also succeeded in designing specific primers for their organism of choice (25-27). In a study by Withers *et al*, a similar genome comparison method was performed on *Pseudoperonospora cubensis* and *Pseudoperonospora humuli* (27). The comparison was first performed *in silico* and subsequently *in vitro.* Using this approach, they were able to identify and determine a large number of specific markers for their organism of interest (27). However, we were not able to perform a similar *in silico* approach here because at the time of data analysis and the preparation of this manuscript, only one *M. mycetomatis* isolate has ever been sequenced and none of *M. fahalii, M. tropicana* and *M. pseudomycetomatis* has ever been sequenced.

In conclusion, since cross-reactivity occurs with the current *M. mycetomatis* specific PCR primer pair 26.1a and 28.3a, we have used a comparative genome approach to identify and designed a new *M. mycetomatis* species-specific PCR primers. We now recommend all reference and local laboratories to use the new PCR primers Mmy-Fw and Mmy-Rv to identify *M. mycetomatis* to the species level.

Furthermore, this comparative genome approach may also be used to design markers for other eumycetoma agents and also other fungi that share conserve ITS region within its genus.

Funding

This study was only supported by internal funding.

References

1. Zijlstra EE, van de Sande WWJ, Welsh O, Mahgoub ES, Goodfellow M, Fahal AH. 2016. Mycetoma: a unique neglected tropical disease. The Lancet Infectious Diseases 16:100-112.
2. Fahal AH, Hassan MA. 1992. Mycetoma. Br J Surg 79:1138-41.
3. van de Sande WW, Gorkink R, Simons G, Ott A, Ahmed AO, Verbrugh H, van Belkum A. 2005. Genotyping of Madurella mycetomatis by selective amplification of restriction fragments (amplified fragment length polymorphism) and subtype correlation with geographical origin and lesion size. J Clin Microbiol 43:4349-56.
4. Borman AM, Desnos-Ollivier M, Campbell CK, Bridge PD, Dannaoui E, Johnson EM. 2016. Novel Taxa Associated with Human Fungal Black-Grain Mycetomas: Emarellia grisea gen. nov., sp. nov., and Emarellia paragrisea sp. nov. J Clin Microbiol 54:1738-45.
5. Khot PD, Fredricks DN. 2009. PCR-based diagnosis of human fungal infections. Expert Rev Anti Infect Ther 7:1201-21.
6. Kourkoumpetis TK, Fuchs BB, Coleman JJ, Desalermos A, Mylonakis E. 2012. Polymerase Chain Reaction–Based Assays for the Diagnosis of Invasive Fungal Infections. Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America 54:1322-1331.
7. Lass-Flörl C, Mutschlechner W, Aigner M, Grif K, Marth C, Girschikofsky M, Grander W, Greil R, Russ G, Cerkl P, Eller M, Kropshofer G, Eschertzhuber S, Kathrein H, Schmid S, Beer R, Lorenz I, Theurl I, Nachbaur D. 2013. Utility of PCR in Diagnosis of Invasive Fungal Infections: Real-Life Data from a Multicenter Study. Journal of Clinical Microbiology 51:863.
8. Buchheidt D, Hummel M. 2005. Aspergillus polymerase chain reaction (PCR) diagnosis. Med Mycol 43 Suppl 1:S139-45.
9. Yeo SF, Wong B. 2002. Current status of nonculture methods for diagnosis of invasive fungal infections. Clin Microbiol Rev 15:465-84.
10. Bursle E, Robson J. 2016. Non-culture methods for detecting infection. Aust Prescr 39:171-175.

11. Laupland KB, Valiquette L. 2013. The changing culture of the microbiology laboratory. Can J Infect Dis Med Microbiol 24:125-8.
12. Ahmed AO, van Leeuwen W, Fahal A, van de Sande W, Verbrugh H, van Belkum A. 2004. Mycetoma caused by Madurella mycetomatis: a neglected infectious burden. Lancet Infect Dis 4:566-74.
13. WHO. 2016. Sixty-ninth world health assembly WHA69.21 - Addressing the burden of mycetoma. http://www.who.int/neglected_diseases/mediacentre/WHA_69.21_Eng.pdf?ua=1. Accessed
14. Ahmed AO, Mukhtar MM, Kools-Sijmons M, Fahal AH, de Hoog S, van den Ende BG, Zijlstra EE, Verbrugh H, Abugroun ES, Elhassan AM, van Belkum A. 1999. Development of a species-specific PCR-restriction fragment length polymorphism analysis procedure for identification of Madurella mycetomatis. J Clin Microbiol 37:3175-8.
15. Nyuykonge B, Klaassen CHW, Zandijk WHA, de Hoog GS, Ahmed SA, Desnos-Ollivier M, Verbon A, Bonifaz A, van de Sande WWJ. 2020. Diagnostic Implications of Mycetoma Derived from Madurella pseudomycetomatis Isolates from Mexico. Journal of the European Academy of Dermatology and Venereology n/a.
16. Ahmed SA, van den Ende BH, Fahal AH, van de Sande WW, de Hoog GS. 2014. Rapid identification of black grain eumycetoma causative agents using rolling circle amplification. PLoS Negl Trop Dis 8:e3368.
17. van de Sande WW. 2013. Global burden of human mycetoma: a systematic review and meta-analysis. PLoS Negl Trop Dis 7:e2550.
18. de Hoog GS, van Diepeningen AD, Mahgoub el S, van de Sande WW. 2012. New species of Madurella, causative agents of black-grain mycetoma. J Clin Microbiol 50:988-94.
19. Yan J, Deng J, Zhou C-J, Zhong B-Y, Hao F. 2010. Phenotypic and molecular characterization of Madurella pseudomycetomatis sp. nov., a novel opportunistic fungus possibly causing black-grain mycetoma. Journal of clinical microbiology 48:251-257.
20. Smit S, Derks MF, Bervoets S, Fahal A, van Leeuwen W, van Belkum A, van de Sande WW. 2016. Genome Sequence of Madurella mycetomatis mm55, Isolated from a Human Mycetoma Case in Sudan. Genome Announc 4.
21. Ahmed A, Adelmann D, Fahal A, Verbrugh H, van Belkum A, de Hoog S. 2002. Environmental occurrence of Madurella mycetomatis, the major agent of human eumycetoma in Sudan. J Clin Microbiol 40:1031-6.
22. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-402.
23. Christensen AC. 2014. Genes and junk in plant mitochondria-repair mechanisms and selection. Genome biology and evolution 6:1448-1453.
24. Studer Romain A, Dessailly Benoit H, Orengo Christine A. 2013. Residue mutations and their impact on protein structure and function: detecting beneficial and pathogenic changes. Biochemical Journal 449:581-594.
25. Feau N, Beauseigle S, Bergeron MJ, Bilodeau GJ, Birol I, Cervantes-Arango S, Dhillon B, Dale AL, Herath P, Jones SJM, Lamarche J, Ojeda DI, Sakalidis ML, Taylor G, Tsui CKM,

Uzunovic A, Yueh H, Tanguay P, Hamelin RC. 2018. Genome-Enhanced Detection and Identification (GEDI) of plant pathogens. PeerJ 6:e4392.

26. Garcia-Rubio R, Monzon S, Alcazar-Fuoli L, Cuesta I, Mellado E. 2018. Genome-Wide Comparative Analysis of Aspergillus fumigatus Strains: The Reference Genome as a Matter of Concern. Genes 9:363.

27. Withers S, Gongora-Castillo E, Gent D, Thomas A, Ojiambo PS, Quesada-Ocampo LM. 2016. Using Next-Generation Sequencing to Develop Molecular Diagnostics for Pseudoperonospora cubensis, the Cucurbit Downy Mildew Pathogen. Phytopathology 106:1105-1116.

Table 1. The sixteen predicted protein sequences with their corresponding size, primer sequences and annealing temperatures.

Primer set	Sequence length (bp)	E value	Primers (5'-3')		Annealing temperature °C
1	972	740	F	ATGCCTGCCCGGTCAGTTCG	55
			R	CTAGTACATGCCCACAACCG	
2	832	96	F	ATGCGCTTTCTCTCCCTTAC	55
			R	TCAGCACTCCCTGATCAACC	
3	808	35	F	ATGCTGCTCGAAAGGGTGTC	55
			R	TCAACCCCGCCCCGTACCCG	
4	639	0.006	F	ATGCACTTCTTCAACACTGT	55
			R	CTAGACGGAGACACCTAGGG	
5	636	1.3	F	ATGAAGCTCACTGTCTCCCT	55
			R	TCAAAGAACAAAAGAGGCAG	
6	621	1.9	F	ATGAAGTACTCTAGCACTCT	55
			R	TTAGGCCGCCTGGGTGGCCG	
7	564	-	F	ATGAAGCTCATCTCCATCGT	55
			R	TCACAAGAGGTACACAACAG	
8	561	0.28	F	ATGCAGCTCTCGATCGCCAA	55
			R	TTAAAGCAACATAGCCGCGT	
9	677	2.1	F	ATGGATCGCCTCGTCAAACC	55
			R	CTAAGTCAACAGAACGACAG	
10	639	2.5	F	ATGAGGTGGCTCGAGACGAC	55
			R	CTATGGTTGTCCACACCCAT	
Mmy-Fw Mmy-Rv	474	20	F	TCTCCTGTCCTACGACATCTGTGG	59
			R	TTCCTCACCTCCCAGCCCTTT	
12	1089	0.007	F	ATGGTGGAGCAGCTCTTGGT	55
			R	TCAAGGAATCGTTCTCGTAA	
13	852	22	F	ATGCATCAACGACATCTTGC	55
			R	CTAGAATTCCTGACGAGAAA	
14	504	42	F	ATGAAATTCACGGACTCTGG	55
			R	CTACATCAGCGGGCACTCCT	
15	544	5.8	F	ATGACAATCACAATCACAAT	55
			R	AAGCTGGCCCCCGATCACAG	
16	544	0.91	F	AGTAATCTAGTCACAATGGC	55
			R	TCAACCCGTGAAAATATTGC	
*26.1A *28.3A	420	-	F	AATGAGTTGGGCTTTAACGG	58
			R	TCCCGGTAGTGTAGTGTCCCT	
*26.1B *28.3B	360	-	F	GCAACACGCCCTGGGCGA	58
			R	TCCGCGGGGCGTCCGCCGGA	

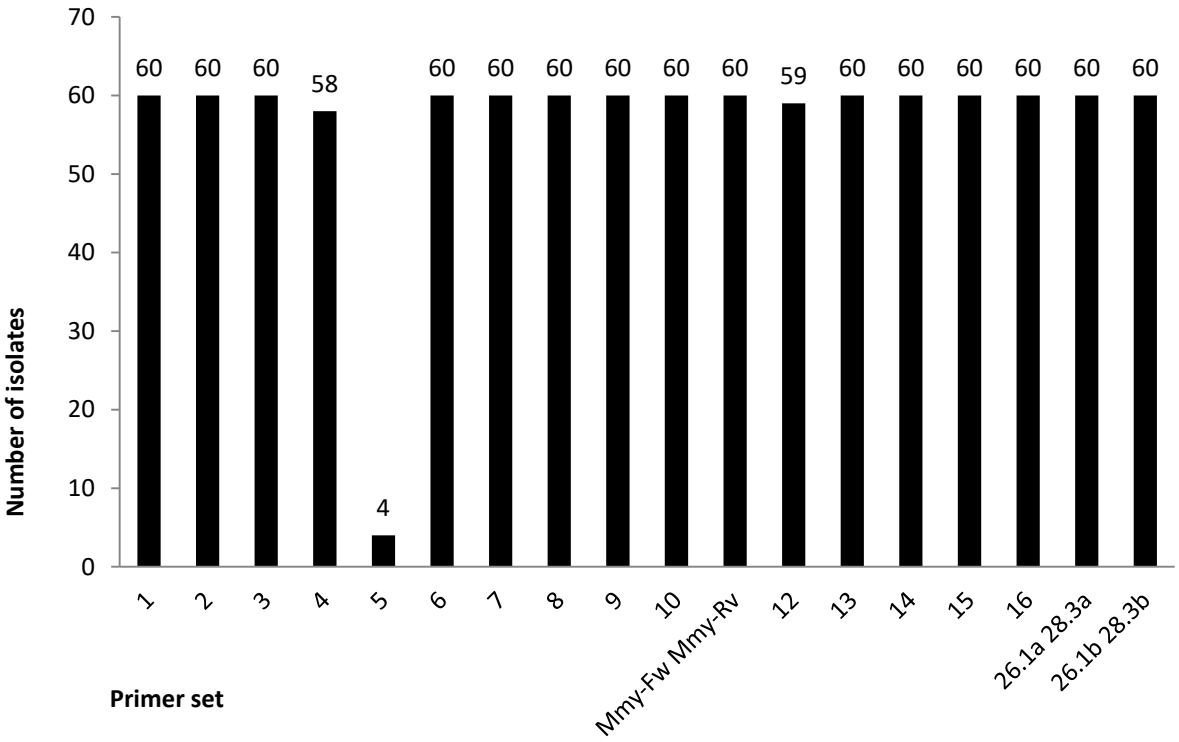
F = forward primer; R = reverse primer; **M. mycetomatis* specific primers designed in 1999 (14).

Table 2. Presence or absence of PCR amplicons of the sixteen primer sets and PCR primers developed in 1999 (14) in the other eumycetoma causing agents and close relatives of *M. mycetomatis.* The grey box highlights the absence of amplicons in all species tested here using Mmy-Fw and Mmy-Rv. Only PCR with bands of the same sizes to *M. mycetomatis* are considered specific to *M. mycetomatis*.

Primer set	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Mmy-Fw Mmy-Rv**	**12**	**13**	**14**	**15**	**16**	*** 26.1A 28.3A**	*** 26.1B 28.3B**
Madurella Tropicana (4)	A	A	C	A	C	A	C	A	A	A	C	A	C	B	B	C	C	A
Madurella fahalii (3)	C	A	C	A	C	C	C	A	A	A	C	C	C	B	B	B	C	A
Madurella pseudomycetomatis (3)	A	A	A	A	C	B	C	A	A	B	C	B	B	C	B	A	B	A
Aspergillus fumigatus (1)	-	-	-	-	-	-	C	-	-	-	C	-	-	-	-	-	-	-
Aspergillus terreus (1)	-	-	-	-	-	-	C	-	-	-	C	-	-	-	-	-	-	-
Chaetomium globosum (2)	-	-	-	-	-	-	B	-	-	-	C	-	-	-	-	-	-	-
Falciformispora senegalensis (4)	B	B	B	B	B	B	C	B	B	B	C	C	B	B	C	B	C	C
Fusarium solani (1)	C	C	B	C	B	B	C	B	B	A	C	C	B	B	B	B	C	C
Medicopsis romeroi (3)	-	-	-	-	-	-	A	-	-	-	C	-	-	-	-	-	C	C
Thielavia subthermophilia (3)	B	C	B	C	C	B	C	B	B	B	C	B	B	B	B	C	C	C
Thielavia terrestris (3)	B	B	B	C	C	B	C	B	B	B	C	B	B	B	B	C	C	C
Trematospheria grisea (4)	-	-	-	-	-	-	C	-	-	-	C	-	-	-	-	-	C	C
Trichophyton rubrum (1)	-	-	-	-	-	-	C	-	-	-	C	-	-	-	-	-	C	C

A: PCR band of the same size; B: PCR band of another size; C: no PCR band. **M. mycetomatis* specific primers designed in 1999 (14).

282 Figure 1. Presence of the 16 PCR amplicons in 60 *M. mycetomatis* isolates tested. Most PCR reactions
283 resulted in amplification in all isolates tested except PCR 4, 5 and 12.
284



285

286