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

Wilson Lim\(^1\), Emmanuel Siddig\(^{1,2}\), Kimberly Eadie\(^1\), Bertrand Nyuykonge\(^1\), Sarah Ahmed\(^{3,\#}\), Ahmed 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, \(^4\) Wageningen University & Research, Department of Plant Science, Wageningen, The Netherlands

\(^\#\) Current address: Center of Expertise in Mycology of Radboud University Medical Center / Canisius Wilhelmina Hospital, Nijmegen, The Netherlands
\(^*\) Current address: Faculty of medical laboratory sciences, University of Khartoum, Sudan

\(*\) w.vandesande@erasmusmc.nl

Abstract

Background

Eumycetoma is a neglected tropical disease most commonly caused by the fungus *Madurella mycetomatis*. Identification of eumycetoma causative agents can only be reliably performed by molecular identification, most commonly by species-specific PCR. The current *M. mycetomatis* specific PCR primers were recently discovered to cross-react with *Madurella pseudomyctomatis*. Here, we used a comparative genome approach to develop a new *M. mycetomatis* specific PCR for species identification.

Methodology

Predicted-protein coding sequences unique to *M. mycetomatis* were first identified in BLASTCLUST based on E-value, size and presence of orthologues. Primers were then developed for 16 unique sequences and evaluated against 60 *M. mycetomatis* isolates and other eumycetoma causing agents including the *Madurella* sibling species. Out of the 16, only one was found to be specific to *M. mycetomatis*.

Conclusion

We have discovered a predicted-protein coding sequence unique to *M. mycetomatis* and have developed a new species-specific PCR to be used as a novel diagnostic marker for *M. mycetomatis*.
Author summary

Mycetoma is a neglected tropical disease characterised by tumorous swellings and grain formation. This disease can be caused by more than 70 different micro-organisms and is categorised into actinomycetoma (caused by bacteria) and eumycetoma (caused by fungi). The most common causative agent of mycetoma is the fungus Madurella mycetomatis. Diagnosis of eumycetoma is often only done clinically or by histopathological examination and culturing of the grains. Unfortunately, that often leads to misidentifications. Molecular identification is currently the most reliable method to identify the causative agents. However, we have recently discovered that the only M. mycetomatis species-specific PCR primers cross-reacts to Madurella pseudomycetomatis. Since all Madurella species cause eumycetoma and have different susceptibilities to antifungal agents, it is important to be able to accurately identify them to the species level. Here we have used a comparative genome approach to identify and design new M. mycetomatis species-specific PCR primers. These primers can be used to identify M. mycetomatis directly from grains and do not cross-react with any of the other eumycetoma causative agents tested. We, therefore, recommend the use of these primers in reference centres and local laboratories to identify M. mycetomatis to the species level.

Introduction

The neglected tropical disease mycetoma presents itself as a subcutaneous chronic granulomatous inflammatory disease and is characterized by tumorous lesions and grain formation [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). Treatment is dependent on the causative agent.

Diagnosis of mycetoma is often only made clinically in endemic areas due to the scarcity of facilities, expertise and financial capacity. Eumycetoma and actinomycetoma can be easily distinguished from each other by the texture and colour of the grains. However, species identification based on the texture and colour of the grain is not possible as many fungal species can produce similar-looking black-grains. Therefore, identification of causative agents is usually done by histopathological examination and culturing of grains. Unfortunately, that often leads to misidentifications [3]. Currently, molecular identification is the most reliable method to identify eumycetoma causative agents down to the species level and the most commonly used technique is amplifying the Internally Transcribed Spacer (ITS) region and sequencing [4]. However, many endemic areas lack the ability to perform DNA sequencing.

In 1999, a species-specific PCR primer based on the internal transcribed spacer (ITS) was developed for Madurella mycetomatis, the most common causative agent of mycetoma [5]. This PCR is currently used at the Mycetoma Reference Center in Khartoum, Sudan. It is performed on DNA obtained from cultured clinical material or directly from grains obtained from patients. Unfortunately, this M. mycetomatis specific PCR primer pair 26.1a and 28.3a was only recently discovered to cross-react with Madurella pseudomycetomatis [6]. Madurella pseudomycetomatis was not yet described at the time when the M. mycetomatis specific PCR was developed [7]. All four Madurella species (Madurella fahalii, Madurella tropicana, M. mycetomatis and M. pseudomycetomatis) are known to cause black grain eumycetoma and have different susceptibilities to antifungal agents. For instance, M. fahalii is not inhibited by itraconazole in vitro, which could have consequences for treatment strategy [8]. This makes identification of causative agents to the species level a must. Since all four Madurella species
share a very conserved ITS region, this has made designing PCR primers specific for \textit{M. mycetomatis} based on that region difficult \cite{8,9}. To circumvent this difficulty, we took a comparative genome approach to design a new set of specific primers for the identification of \textit{M. mycetomatis} by PCR.

Here, we report a new set of diagnostic DNA primers for \textit{M. mycetomatis} identified from the genome of \textit{M. mycetomatis} \cite{10}.

\section*{Materials and methods}

\subsection*{Ethics statement}

This study was approved by the Mycetoma Research Center, Khartoum, Sudan (IRB, No. 11/2018). Written informed consent was obtained from each adult patient, and assent was taken from minors (aged below 18 years) with written consent from their guardian.

\subsection*{Fungal isolates, patient grains and DNA isolation}

A total of 95 fungal isolates were used in this study; 60 \textit{M. mycetomatis}, 4 \textit{M. tropicana}, 3 \textit{M. fahalii}, 3 \textit{M. pseudomykctomatis}, 1 \textit{Aspergillus fumigatus}, 1 \textit{Aspergillus terreus}, 2 \textit{Chaetomium globosum}, 4 \textit{Falciformispora senegalensis}, 1 \textit{Fusarium solani}, 3 \textit{Medicopsis romeroi}, 2 \textit{Scedosporium apiospermum}, 3 \textit{Thielavia terrestris}, 3 \textit{Thielavia subthermophilia}, 4 \textit{Trematospheria grisea} and 1 \textit{Trichophyton rubrum}. Most 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 \textit{M. mycetomatis} isolates originated from mycetoma patients. Isolates are maintained on Sabouraud Dextrose (SAB) agar at either 37˚C or room temperature depending on the fungal species. Black grains were obtained from a total of 16 eumycetoma infected patients seen at the Mycetoma Research Center in Sudan. Nine patients were confirmed to be infected with \textit{M. mycetomatis}, four with \textit{F. senegalensis}, and three with \textit{F. tompkinsii}. DNA from fungal isolates and grains were isolated as described earlier using the ZR Fungal/Bacterial DNA MicroPrep kit (Zymo Research, USA) \cite{11}. All isolates were identified to the species level based on morphology, polymerase chain reaction (PCR)-based restriction fragment length polymorphisms, and sequencing of the ITS regions \cite{5,12,13}.

\subsection*{Identifying specific predicted protein-coding sequences to \textit{M. mycetomatis}}

Predicted protein-coding sequences (PPCS) of \textit{M. mycetomatis} were obtained from the published genome sequence of \textit{M. mycetomatis} isolate mm55, accession number LCTW0000000, BioProject PRJNA267680 \cite{10}. To determine their specificity to \textit{M. mycetomatis}, a bioinformatical comparison of these sequences to the genome of other organisms was performed using BLASTCLUST \cite{14}. The specificity of these PPCS were determined based on presence of orthologues, E-value and fragment size. Orthologues were defined as sequences with greater than 85% amino acid similarity to the tested \textit{M. mycetomatis} PPCS. \textit{M. mycetomatis} PPCS with no orthologues present in the genomes of other organisms, E- value of 0.003 and higher and size between 400 bp and 1100 bp were considered to be specific to \textit{M. mycetomatis} and were chosen for further analysis.

\subsection*{Primer design and PCR conditions}

Forward and reverse PCR primers were designed according to the nucleotide sequence of the PPCS of interest. Primer sequences are depicted in Table 1. PCR reaction contained 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. PCR conditions were as follows: 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 in 2% agarose gel (Sphaero Q).

### Results and discussion

Since we have demonstrated that the currently used *M. mycetomatis* specific PCR cross-reacted with *M. pseudomykctomatis*, there was a need to develop a new *M. mycetomatis* specific PCR for

---

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: ATGCCCTGCCGCTCAGTTCG | 55             |
|            |                     |         | R: CTAGTACATGCCACAAAAGCG |               |
| 2          | 832                 | 96      | F: ATGGCTTCTCCCCCCTTAC | 55             |
|            |                     |         | R: TCAAGCCTCTCTGATCAACCC |               |
| 3          | 808                 | 35      | F: ATGHCTGCTGAGACGGGGTTGTC | 55             |
|            |                     |         | R: TCAACCCTCCCCCTGATACCCG |               |
| 4          | 639                 | 0.006   | F: ATGCACCTCTTTCAACACTG | 55             |
|            |                     |         | R: CTAGACGGAGACACCTAGGG |               |
| 5          | 636                 | 1.3     | F: ATGAAGCTACTGTCTCCCT | 55             |
|            |                     |         | R: TCAAAGAAGAAAAAGGCAGA |               |
| 6          | 621                 | 1.9     | F: ATGAAGACTACTGACTCT | 55             |
|            |                     |         | R: TTAGGCGCTGAGGCCGCT |               |
| 7          | 564                 | -       | F: ATGAAGCTACATCCATG | 55             |
|            |                     |         | R: TCAACAGGAGTGACAAACAG |               |
| 8          | 561                 | 0.28    | F: ATGCAGCTCTCGATGCGA | 55             |
|            |                     |         | R: TAAAAGCAACATAAGCCGCT |               |
| 9          | 677                 | 2.1     | F: ATGATGGCGCTGATCAACCC | 55             |
|            |                     |         | R: CTAAGTCAACAGAGACGAC |               |
| 10         | 639                 | 2.5     | F: ATGAGGTGCCGAGACGAC | 55             |
|            |                     |         | R: CTATGTTGTGACACCAACAT |               |
| Mmy-Fw     | 474                 | 20      | F: TCTCTCTGCTCTACAGCATCTGTGG | 59             |
|            |                     |         | R: TCCCTACCTGCCAGGCCCTTT |               |
| Mmy-Rv     | 1089                | 0.007   | F: ATGTTAGGACGCTCTTGGT | 55             |
|            |                     |         | R: TCAAGAATCTTCTCCTGTA |               |
| 12         | 852                 | 22      | F: ATGCATCAACGACATCTTG | 55             |
|            |                     |         | R: CTAAGATCTCTGACGAAGAA |               |
| 13         | 504                 | 42      | F: ATGAAATTCAGCAGACTCTTG | 55             |
|            |                     |         | R: CTAACATCAGCGGCACTCCT |               |
| 14         | 544                 | 5.8     | F: ATGCAATACCAATCACCAT | 55             |
|            |                     |         | R: AAGCTGGCCCGATACACAG |               |
| 15         | 544                 | 0.91    | F: ATGAATCTAGTCACATG | 55             |
|            |                     |         | R: TCAACCCGTGAAAATATTTG |               |
| *26.1A     | 420                 | -       | F: AATGAGTTGGCCITTAAACGG | 58             |
| *28.3A     |                     |         | R: TCCCGGTAGTGTAGTGTCC |               |
| *26.1B     | 360                 | -       | F: GCAACAGGCCCCGAGGCGA | 58             |
| *28.3B     |                     |         | R: TCCGGGGGCGCTCCGCGGA |               |

*M. mycetomatis* specific primers designed in 1999 [5].

https://doi.org/10.1371/journal.pntd.0008897.t001
proper species identification. From the genome of *M. mycetomatis*, 350 predicted protein-coding sequences (PPCS) were randomly selected and analysed. We chose to analyse PPCS because these protein-coding sequences are likely to be more stable than non-coding sequences [15,16]. To ensure that they can be easily amplified through PCR, we preferentially chose PPCS with sizes between 400 and 1100 bp. From the initial 350 PPCS, the top 16 candidates that fitted our requirement based on specificity and size were chosen for PCR development.

PCR primers for the 16 candidates were then designed (Table 1). To ascertain that these primers would amplify their targets in all *M. mycetomatis* isolates, they were evaluated in 60 *M. mycetomatis* isolates from different geographical origins, genotypic backgrounds and phenotypic appearance. Out of the 16 primer sets tested, 13 were positive in all 60 *M. mycetomatis* isolates tested (Fig 1). Primer sets 4, 5 and 12 were present in 58, 4 and 59 isolates, respectively (Fig 1). To determine the specificity of the 13 positive primer sets to *M. mycetomatis*, 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. The amplicon generated by Mmy-Fw and Mmy-Rv appears to be a putative single-copy gene. To determine if these PCR primers were as sensitive as the currently used ones, we compared the two PCRs head-on. Mmy-Fw and Mmy-Rv were able to detect DNA concentrations as low as 5 pg. This is only slightly less sensitive compared to the currently used diagnostic PCR primer pair 26.1a and 28.3a that is able to detect DNA at 0.5 pg.

Primers Mmy-Fw and Mmy-Rv were also tested on DNA extracted from grains obtained from eumycetoma patients. As shown in Fig 2, amplicons were only observed when DNA

---

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

https://doi.org/10.1371/journal.pntd.0008897.g001
from *M. mycetomatis* grains were present. The primers did not cross-react to DNA obtained from *F. senegalensis* or *F. tompkinsii* grains (Fig 2). Our findings show that these primers are sufficiently sensitive to be used in diagnosis directly from clinical specimens.

One of the advantages of using this comparative genome approach is that primer designs are less constrained since the targeted genes are unique. With this method, we were able to design primers that can distinguish between *M. mycetomatis* and *M. pseudomycetomatis*.

Other studies have also succeeded in designing specific primers for their organism of choice.
using this approach [17–19]. In a study by Withers et al, a similar genome comparison method was performed on *Pseudoperonospora cubensis* and *Pseudoperonospora humuli* [19]. 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 while reducing the number of diagnostic candidates to validate with PCR [19]. However, a similar *in silico* approach could not be performed in our study because at the time of data analysis and the preparation of this manuscript, only the genome of one *M. mycetomatis* isolate and none of *M. fahalii*, *M. tropicana* and *M. pseudomycetomatis* was 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 new *M. mycetomatis* species-specific PCR primers. Since new fungi causing eumycetoma are still being discovered, proper identification of its causative agents can help to fully understand the epidemiology and global burden of this disease. Thus, there is clearly a need for a specific PCR marker to identify its causative agents. We recommend reference centers such as the WHO collaborative Mycetoma Reference Center in Khartoum, Sudan and yet to be established reference laboratories in other endemic countries 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.

**Author Contributions**

**Conceptualization:** Wendy WJ van de Sande.

**Data curation:** Wilson Lim, Sandra Smit, Wendy WJ van de Sande.

**Formal analysis:** Wilson Lim.

**Funding acquisition:** Wendy WJ van de Sande.

**Investigation:** Wilson Lim, Kimberly Eadie.

**Methodology:** Wilson Lim, Emmanuel Siddig, Kimberly Eadie, Bertrand Nyuykonge.

**Project administration:** Wendy WJ van de Sande.

**Resources:** Wilson Lim, Sarah Ahmed, Ahmed Fahal, Sandra Smit, Wendy WJ van de Sande.

**Supervision:** Annelies Verbon, Wendy WJ van de Sande.

**Validation:** Wilson Lim.

**Visualization:** Wilson Lim.

**Writing – original draft:** Wilson Lim.

**Writing – review & editing:** Wilson Lim, Bertrand Nyuykonge, Sarah Ahmed, Ahmed Fahal, Annelies Verbon, Sandra Smit, Wendy WJ van de Sande.

**References**

1. Zijlstra EE, van de Sande WWJ, Welsh O, Mahgoub ES, Goodfellow M, Fahal AH. Mycetoma: a unique neglected tropical disease. *The Lancet Infectious Diseases*. 2016; 16(1):100–12. https://doi.org/10.1016/S1473-3099(15)00359-X PMID: 26738840

2. Fahal AH, Hassan MA. Mycetoma. *Br J Surg.* 1992; 79(11):1138–41. https://doi.org/10.1002/bjs.1800791107 PMID: 14678883

3. Borman AM, Desnos-Ollivier M, Campbell CK, Bridge PD, Dannaoui E, Johnson EM. Novel Taxa Associated with Human Fungal Black-Grain Mycetomas: *Emarella grisea* gen. nov., sp. nov., and *Emarella*...
paragrisea sp. nov. Journal of clinical microbiology. 2016; 54(7):1738–45. https://doi.org/10.1128/JCM.00477-16 PMID: 27076666

4. Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, et al. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proceedings of the National Academy of Sciences. 2012; 109(16):6241. https://doi.org/10.1073/pnas.1117018109 PMID: 22454494

5. Ahmed AO, Mukhtar MM, Kools-Sijmons M, Fahal AH, de Hoog S, van den Ende BG, et al. Development of a species-specific PCR-restriction fragment length polymorphism analysis procedure for identification of Madurella mycetomatis. Journal of clinical microbiology. 1999; 37(10):1185–86. https://doi.org/10.1128/JCM.37.10.3175-3178.1999 PMID: 10488173

6. Nyuykonge B, Klaassen CHW, Zandijk WHA, de Hoog GS, Ahmed SA, Desnos-Ollivier M, et al. Diagnostic Implications of Mycetoma Derived from Madurella pseudomycetomatis Isolates from Mexico. Journal of the European Academy of Dermatology and Venereology. 2020; n/a(n/a). https://doi.org/10.1111/jdv.16402 PMID: 32233084

7. Ahmed SA, van den Ende BH, Fahal AH, van de Sande WW, de Hoog GS. Rapid identification of black grain eumycetoma causative agents using rolling circle amplification. PLoS Negl Trop Dis. 2014; 8(12):e3368. https://doi.org/10.1371/journal.pntd.0003368 PMID: 25474355

8. de Hoog GS, van Diepening AD, Mahgoub el S, van de Sande WW. New species of Madurella, causative agents of black-grain mycetoma. Journal of clinical microbiology. 2012; 50(3):988–94. https://doi.org/10.1128/JCM.05477-11 PMID: 22205798

9. Yan J, Deng J, Zhou C-J, Zhong B-Y, Hao F. Phenotypic and molecular characterization of Madurella pseudomycetomatis sp. nov., a novel opportunistic fungus possibly causing black-grain mycetoma. Journal of clinical microbiology. 2010; 48(1):251–7. https://doi.org/10.1128/JCM.00018-09 PMID: 19923486

10. Smit S, Derks MF, Bervoets S, Fahal A, van Leeuwen W, van Belkum A, et al. Genome Sequence of Madurella mycetomatis mm55, Isolated from a Human Mycetoma Case in Sudan. Genome Announc. 2016; 4(3). https://doi.org/10.1128/genomeA.00418-16 PMID: 27231361

11. Lim W, Eadie K, Horst-Kreft D, Ahmed SA, Fahal AH, van de Sande WWJ. VNTR confirms the heterogeneity of Madurella mycetomatis and is a promising typing tool for this mycetoma causing agent. Medical Mycology. 2019; 57(4):434–40. https://doi.org/10.1093/mmy/myy055 PMID: 30086253

12. Ahmed A, Adelmann D, Fahal A, Verbrugh H, van Belkum A, de Hoog S. Environmental occurrence of Madurella mycetomatis, the major agent of human eumycetoma in Sudan. Journal of clinical microbiology. 2002; 40(3):1031–6. https://doi.org/10.1128/jcm.40.3.1031-1036.2002 PMID: 11890433

13. van de Sande WW, Gorkink R, Simons G, Ott A, Ahmed AO, Verbrugh H, et al. Genotyping of Madurella mycetomatis by selective amplification of restriction fragments (amplified fragment length polymorphism) and subtype correlation with geographical origin and lesion size. Journal of clinical microbiology. 2005; 43(9):4349–56. https://doi.org/10.1128/JCM.43.9.4349-4356.2005 PMID: 16145076

14. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997; 25(17):3389–402. https://doi.org/10.1093/nar/25.17.3389 PMID: 9254694

15. Christensen AC. Genes and junk in plant mitochondria-repair mechanisms and selection. Genome Biol Evol. 2014; 6(6):1448–53. https://doi.org/10.1093/gbe/evu115 PMID: 24904012

16. Studer Romain A, Dessailly Benoit H, Oreno Christine A. Residue mutations and their impact on protein structure and function: detecting beneficial and pathogenic changes. Biochemical Journal. 2013; 449(3):581–94. https://doi.org/10.1042/BJ20121221 PMID: 23301657

17. Feau N, Beauseigle S, Bergeron MJ, Bilodeau GJ, Birol I, Cervantes-Aranzo S, et al. Genome-Enhanced Detection and Identification (GEDI) of plant pathogens. PeerJ. 2018; 6:e4392. https://doi.org/10.7717/peerj.4392 PMID: 29492338

18. Garcia-Rubio R, Monzon S, Alcazar-Fuoli L, Cuesta I, Mellado E. Genome-Wide Comparative Analysis of Aspergillus fumigatus Strains: The Reference Genome as a Matter of Concern. Genes (Basel). 2018; 9(7):363. https://doi.org/10.3390/genes9070363 PMID: 30029559

19. Withers S, Gongora-Castillo E, Gent D, Thomas A, Ojiamb o PS, Quesada-Ocampo LM. Using Next-Generation Sequencing to Develop Molecular Diagnostics for Pseudoperonospora cubensis, the Cucurbit Downy Mildew Pathogen. Phytopathology. 2016; 106(10):1105–16. https://doi.org/10.1094/PHYTO-10-15-0260-FI PMID: 27314624