


RESEARCH ARTICLE

Open Access



Decision criteria for MALDI-TOF MS-based identification of filamentous fungi using commercial and in-house reference databases

Anne-Cécile Normand^{1*} , Carole Cassagne¹, Magali Gautier¹, Pierre Becker², Stéphane Ranque¹, Marijke Hendrickx² and Renaud Piarroux¹

Abstract

Background: Several Matrix-Assisted Laser Desorption/Ionization Time-of-Flight mass spectrometry protocols, which differ in identification criteria, have been developed for mold and dermatophyte identification. Currently, the most widely used approach is Bruker technology, although no consensus concerning the log(score) threshold has been established. Furthermore, it remains unknown how far increasing the number of spots to compare results might improve identification performance.

In this study, we used in-house and Bruker reference databases as well as a panel of 422 isolates belonging to 126 species to test various thresholds. Ten distinct identification algorithms requiring one to four spots were tested.

Results: Our findings indicate that optimal results were obtained by applying a decisional algorithm in which only the highest score of four spots was taken into account with a 1.7 log(score) threshold. Testing the entire panel enabled identification of 87.41% (in-house database) and 35.15% (Bruker database) of isolates, with a positive predictive value (PPV) of 1 at the genus level for both databases as well as 0.89 PPV (in-house database) and 0.72 PPV (Bruker database) at the species level. Applying the same rules to the isolates for which the species were represented by at least three strains in the database enabled identification of 92.1% (in-house database) and 46.6% (Bruker database) of isolates, with 1 PPV at the genus level for both databases as well as 0.95 PPV (in-house database) and 0.93 PPV (Bruker database) at the species level.

Conclusions: Depositing four spots per extract and lowering the threshold to 1.7, a threshold which is notably lower than that recommended for bacterial identification, decreased the number of unidentified specimens without altering the reliability of the accepted results. Nevertheless, regardless of the criteria used for mold and dermatophyte identification, commercial databases require optimization.

Keywords: Mass Spectrometry, Matrix-Assisted Laser Desorption/Ionization, Filamentous Fungi, Interpretation criteria, log(score) threshold

* Correspondence: anne-cecile.normand@ap-hm.fr

¹Laboratoire de Parasitologie-Mycologie, CHU Timone, Université de la Méditerranée, Marseille, France

Full list of author information is available at the end of the article



Background

Until a few years ago, medical laboratories performed mold identification using phenotypic characterization of colonies and occasional sequencing of informative DNA targets. Currently, the MALDI-TOF MS (Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry) approach appears to be a promising method to identify molds and dermatophytes [1–8].

Many teams around the world have demonstrated that filamentous fungi could be identified by using MALDI-TOF MS. They first focused attention on a few specific mold and dermatophyte genera [1, 9–16]. Only after 2010 did research teams begin optimizing protocols to establish a MALDI-TOF MS-based identification approach for routine diagnosis of filamentous fungi and dermatophytes [2–7, 17–21]. Currently, four commercial systems are available for routine diagnosis, and an increasing number of publications have tested their performance in identifying molds and dermatophytes. Some teams have assessed the Saramis [12, 18, 22, 23], Vitek MS [24–28] or Andromas [29] systems; however, most publications refer to the Bruker technique [1–8, 16, 17, 19–21, 26, 30, 31]. According to the Bruker recommendations, the level of similarity between an unknown tested specimen and a reference sample is indicated by a $\log(\text{score})$, which is referred to herein as “score”. A score > 2.3 indicates “highly probable species identification”, a score > 2 and < 2.299 indicates “secure genus identification, probable species identification”, a score > 1.7 and < 1.999 indicates “probable genus identification”, and a score < 1.7 indicates “unreliable identification”.

However, these score thresholds have been designed for bacteria identification and are not necessarily appropriate for fungi (especially molds and dermatophytes) while using the manufacturer’s reference database [4, 7, 17, 32, 33]. The majority of publications dealing with MALDI-TOF MS routine identification of filamentous fungi have used a single score threshold, usually ranging from 1.7 to 2.0, depending on the study [3, 4, 6, 8, 17]. However, a study by Shultness et al. [8] has compared the 1.7 and 2.0 score thresholds for MALDI-TOF MS identification after a liquid culture growth step in which no assessment was performed to justify these thresholds. Additionally, while the manufacturer typically recommends a single spot for bacterial identification, some authors have advocated the use of several spots of the same sample to identify molds and dermatophytes [2, 21]. According to these authors, the concordance between identification results can be tested using replicates. Theoretically, this may improve the reliability of identification results, although no study has thoroughly evaluated the optimal number of spots and $\log(\text{score})$ threshold for mold identification. Altogether, the use of various thresholds and validation

criteria for mold and dermatophyte identification complicate the interpretation of results.

The present study assessed 10 distinct identification algorithms combining several thresholds for one to four spots per sample. Because the Bruker reference database is not comprehensive enough to enable proper identification of many mold species, [4, 8] identification algorithms were assessed using an extensive in-house database (5044 references belonging to 619 species and 165 genera) and the Bruker commercial database (V3.2.1.1 : 604 references belonging to 237 species and 67 genera). Our objective was to optimize the fungal identification workflow in laboratories equipped with a Bruker MALDI-TOF MS identification platform.

Methods

Reference databases

The in-house reference database was constructed using a Microflex LT system (Bruker Daltonics, Bremen, Germany) in coordination with the BCCM/IHEM culture collection and the Mycology Laboratory of the AP-HM with the protocol described by Normand et al. [6] (several strains per species, four subcultures per strain and 10 spots per subculture, for a total of four references per strain). Details of the MS in-house database are listed in Additional file 1: Table S1.

Assessment panel

The assessment panel consisted of filamentous fungus strains that were identified via both DNA sequencing and MALDI-TOF MS. The strains were either selected at the Mycology Laboratory of Marseille ($n = 224$) or the BCCM/IHEM in Brussels ($n = 198$). Of the 224 strains from Marseille, 177 were included because the identification at the species level had to be either clarified or confirmed at the time of the study, while the remaining 47 were randomly selected among frequently identified species in the clinical laboratory. The 198 Belgian strains either originated from the BCCM/IHEM collection or were clinical strains from a previous study [34]. The 422-strain panel included 126 different species belonging to 38 genera, of which 280 strains were represented three or more times in the MALDI-TOF MS reference database (191 strains for the commercial Bruker database), 130 strains were represented once or twice (76 strains for the commercial Bruker database), and 12 strains were not represented in the reference database (155 strains for the commercial Bruker database) (Tables 1 and 2). All samples were first included in the study panel, even those corresponding to species absent from the databases to ensure that the tested algorithms correctly excluded absent species. Next, the identification process was tested again, this time only including the isolates corresponding to species represented by at least three distinct strains in the

Table 1 Panel description obtained after DNA sequencing of the strains and representation of the entire set of species in the in-house reference spectrum database

Species not represented in the in-house reference database	TOTAL (n = 12)	Collection strains BCCM/IHEM	Clinical strains (Marseille Hospital)	Species represented once or twice in the in-house reference database	TOTAL (n = 130)	Collection strains BCCM/IHEM	Clinical strains (Marseille Hospital)	Species represented three or more times in the in-house reference database	TOTAL (n = 280)	Collection strains BCCM/IHEM	Clinical strains (Marseille Hospital)
<i>Cladosporium bruhnei</i>	1	1	0	<i>Acremonium breve</i>	1	0	1	<i>Alternaria alternata</i>	2	1	1
<i>Cladosporium halotolerans</i>	2	2	0	<i>Acremonium strictum</i>	6	6	0	<i>Aspergillus alabamensis</i>	4	0	4
<i>Colletotrichum lineola</i>	1	1	0	<i>Alternaria infectaria</i>	2	1	1	<i>Aspergillus caesiellus</i>	1	1	0
<i>Trichoderma orientale</i>	1	1	0	<i>Arthrinium arundinis</i>	1	0	1	<i>Aspergillus calidoustus</i>	10	2	8
<i>Mucor hiemalis</i>	1	1	0	<i>Aspergillus dimorphicus</i>	1	0	1	<i>Aspergillus carbonarius</i>	1	0	1
<i>Penicillium atramentosum</i>	1	1	0	<i>Aspergillus fischeri</i>	2	0	2	<i>Aspergillus carneus</i>	1	0	1
<i>Penicillium concentricum</i>	1	1	0	<i>Aspergillus flavipes</i>	3	2	1	<i>Aspergillus chevalieri</i>	4	2	2
<i>Phaeosphaeria avenaria</i>	1	1	0	<i>Aspergillus floccosus</i>	3	0	3	<i>Aspergillus clavatus</i>	2	2	0
<i>Pleospora papaveracea</i>	1	1	0	<i>Aspergillus fumigatiifinis</i>	2	0	2	<i>Aspergillus creber</i>	1	1	0
<i>Trichoderma reesei</i>	2	2	0	<i>Aspergillus iizukae</i>	1	0	1	<i>Aspergillus flavus</i>	15	5	10
				<i>Aspergillus quadrilineatus</i>	1	0	1	<i>Aspergillus fumigatus</i>	20	4	16
				<i>Aspergillus ruber</i>	2	2	0	<i>Aspergillus hiratsukae</i>	1	0	1
				<i>Aspergillus sclerotiorum</i>	3	1	2	<i>Aspergillus hollandicus</i>	5	4	1
				<i>Aspergillus subolivaceus</i>	1	0	1	<i>Aspergillus insuetus</i>	1	0	1
				<i>Aspergillus unguis</i>	4	4	0	<i>Aspergillus nidulans</i>	8	4	4
				<i>Aureobasidium pullulans</i>	3	2	1	<i>Aspergillus niger</i>	4	2	2
				<i>Cladosporium cladosporioides</i>	1	1	0	<i>Aspergillus nomius</i>	1	0	1
				<i>Cladosporium pseudocladosporioides</i>	2	2	0	<i>Aspergillus parasiticus</i>	2	1	1
				<i>Cladosporium sphaerospermum</i>	1	1	0	<i>Aspergillus sydowii</i>	6	2	4
				<i>Cochliobolus hawaiiensis</i>	2	0	2	<i>Aspergillus tamarii</i>	2	2	0
				<i>Curvularia lunata</i>	2	2	0	<i>Aspergillus terreus</i>	10	3	7
				<i>Fusarium dimerum</i>	2	1	1	<i>Aspergillus tubingensis</i>	29	1	28
				<i>Geosmithia argillacea</i>	1	0	1	<i>Aspergillus versicolor</i>	3	2	1

Table 1 Panel description obtained after DNA sequencing of the strains and representation of the entire set of species in the in-house reference spectrum database (Continued)

<i>Geosmithia pallida</i>	1	1	0	0	1	0	1	1
<i>Hexagonia hydnooides</i>	1	0	1	1	3	2	1	1
<i>Lecytophthora</i> sp.	1	1	0	0	2	2	0	0
<i>Macrophomina phaseolina</i>	1	0	1	1	2	2	0	0
<i>Mucor circinelloides</i>	5	5	0	0	2	0	2	2
<i>Penicillium asturianum</i>	1	0	1	1	1	0	1	1
<i>Penicillium brasilianum</i>	3	1	2	2	21	15	6	6
<i>Penicillium cecidicola</i>	2	1	1	1	10	1	9	9
<i>Penicillium citrinum</i>	5	4	1	1	1	0	1	1
<i>Penicillium corylophilum</i>	4	2	2	2	7	6	1	1
<i>Penicillium crateriforme</i>	2	1	1	1	1	1	0	0
<i>Penicillium diversum</i>	1	1	0	0	2	2	0	0
<i>Penicillium expansum</i>	2	2	0	0	2	2	0	0
<i>Penicillium funiculosum</i>	2	0	2	2	1	0	1	1
<i>Penicillium glabrum</i>	9	3	6	6	2	0	2	2
<i>Penicillium helicum</i>	1	0	1	1	3	2	1	1
<i>Talaromyces marneffei</i>	1	0	1	1	3	2	1	1
<i>Penicillium nalgiovense</i>	1	0	1	1	27	13	14	14
<i>Penicillium oxalicum</i>	1	1	0	0	3	2	1	1
<i>Penicillium pinophilum</i>	3	0	3	3	1	1	0	0
<i>Penicillium polonicum</i>	6	6	0	0	2	2	0	0
<i>Penicillium purpurogenum</i>	1	0	1	1	6	5	1	1
<i>Penicillium raistrickii</i>	1	1	0	0	11	2	9	9
<i>Penicillium rolsfii</i>	1	0	1	1	3	2	1	1
<i>Penicillium rugulosum</i>	1	1	0	0	4	1	3	3
<i>Penicillium sizovae</i>	2	0	2	2	15	13	2	2
<i>Penicillium spinulosum</i>	2	0	2	2	2	2	0	0
<i>Penicillium terrigenum</i>	1	0	1	1	1	0	1	1
<i>Penicillium variabile</i>	2	0	2	2	8	0	8	8
	2	0	2	2				

Table 1 Panel description obtained after DNA sequencing of the strains and representation of the entire set of species in the in-house reference spectrum database (Continued)

<i>Phanerochaete</i>					
<i>chrysosporium</i>					
<i>Pithomyces chartarum</i>	3	3	0		
<i>Pseudallescheria boydii</i>	1	0	1		
<i>Pseudallescheria ellipsoidea</i>	1	1	0		
<i>Rhizomucor variabilis</i>	2	0	2		
<i>Rhizopus microsporus</i>	1	1	0		
<i>Scedosporium aurantiacum</i>	6	6	0		
<i>Tyromyces fissilis</i>	3	0	3		

Table 2 Representation of the entire set of species included in the Bruker reference spectrum database

Species not represented in the Bruker reference database	Total (n = 155)	List of the 26 species represented once or twice in the Bruker reference database	Total (n = 76)	List of the 25 species represented three or more times in the Bruker reference database	Total (n = 191)
<i>Acremonium breve</i>	1	<i>Acremonium strictum</i>	6	<i>Alternaria alternata</i>	2
<i>Alternaria infectoria</i>	2	<i>Aspergillus clavatus</i>	2	<i>Aspergillus flavus</i>	15
<i>Arthrinium arundinis</i>	1	<i>Aspergillus hollandicus</i>	5	<i>Aspergillus fumigatus</i>	20
<i>Aspergillus alabamensis</i>	4	<i>Aspergillus sclerotiorum</i>	3	<i>Aspergillus nidulans</i>	8
<i>Aspergillus caesiellus</i>	1	<i>Aspergillus sydowii</i>	6	<i>Aspergillus niger</i>	4
<i>Aspergillus calidoustus</i>	10	<i>Aspergillus tamarii</i>	2	<i>Aspergillus parasiticus</i>	2
<i>Aspergillus carbonarius</i>	1	<i>Aspergillus unguis</i>	4	<i>Aspergillus terreus</i>	10
<i>Aspergillus carneus</i>	1	<i>Beauveria bassiana</i>	3	<i>Aspergillus versicolor</i>	3
<i>Aspergillus chevalieri</i>	4	<i>Cladosporium cladosporioides</i>	1	<i>Aureobasidium pullulans</i>	3
<i>Aspergillus creber</i>	1	<i>Curvularia lunata</i>	2	<i>Exophiala dermatitidis</i>	2
<i>Aspergillus dimorphicus</i>	1	<i>Epicoccum nigrum</i>	2	<i>Fusarium oxysporum</i>	21
<i>Aspergillus fischeri</i>	2	<i>Fusarium dimerum</i>	2	<i>Fusarium proliferatum</i>	10
<i>Aspergillus flavipes</i>	3	<i>Fusarium equiseti</i>	1	<i>Fusarium solani</i>	7
<i>Aspergillus floccosus</i>	3	<i>Fusarium verticillioides</i>	1	<i>Lichtheimia corymbifera</i>	2
<i>Aspergillus fumigatiiformis</i>	2	<i>Galactomyces geotrichum</i>	2	<i>Microsporum canis</i>	2
<i>Aspergillus hiratsukae</i>	1	<i>Mucor circinelloides</i>	5	<i>Purpureocillium lilacinus</i>	3
<i>Aspergillus iizukae</i>	1	<i>Penicillium citrinum</i>	5	<i>Paecilomyces variotii</i>	3
<i>Aspergillus insuetus</i>	1	<i>Penicillium corylophilum</i>	4	<i>Penicillium chrysogenum</i>	27
<i>Aspergillus nomius</i>	1	<i>Penicillium funiculosum</i>	2	<i>Penicillium expansum</i>	2
<i>Aspergillus quadrilineatus</i>	1	<i>Penicillium purpurogenum</i>	1	<i>Penicillium glabrum</i>	9
<i>Aspergillus ruber</i>	2	<i>Penicillium rugulosum</i>	1	<i>Scedosporium apiospermum</i>	11
<i>Aspergillus subolivaceus</i>	1	<i>Rhizopus microsporus</i>	1	<i>Schizophyllum commune</i>	4
<i>Aspergillus tubingensis</i>	29	<i>Rhizopus oryzae</i>	6	<i>Scopulariopsis brevicaulis</i>	15
<i>Aspergillus westerdijkiae</i>	1	<i>Scedosporium aurantiacum</i>	6	<i>Trichophyton interdigitale</i>	1
<i>Cladosporium bruhnei</i>	1	<i>Scedosporium prolificans</i>	3	<i>Trichophyton rubrum</i>	8
<i>Cladosporium halotolerans</i>	2	<i>Trichoderma longibrachiatum</i>	2		
<i>Cladosporium pseudocladosporioides</i>	2				
<i>Cladosporium sphaerospermum</i>	1				
<i>Cochliobolus hawaiiensis</i>	2				
<i>Colletotrichum lineola</i>	1				
<i>Fomitopsis pinicola</i>	2				
<i>Fusarium sacchari</i>	1				
<i>Geosmithia argillacea</i>	1				
<i>Geosmithia pallida</i>	1				
<i>Hexagonia hydroides</i>	1				
<i>Trichoderma orientale</i>	1				
<i>Lecythophora sp.</i>	1				
<i>Macrophomina phaseolina</i>	1				
<i>Microsporum audouinii</i>	1				
<i>Mucor hiemalis</i>	1				
<i>Penicillium asturianum</i>	1				
<i>Penicillium atramentosum</i>	1				

Table 2 Representation of the entire set of species included in the Bruker reference spectrum database (*Continued*)

<i>Penicillium brasilianum</i>	3
<i>Penicillium cecidicola</i>	2
<i>Penicillium concentricum</i>	1
<i>Penicillium crateriforme</i>	2
<i>Penicillium crustosum</i>	3
<i>Penicillium diversum</i>	1
<i>Penicillium griseofulvum</i>	1
<i>Penicillium helicum</i>	1
<i>Talaromyces marnefei</i>	1
<i>Penicillium nalgiovense</i>	1
<i>Penicillium oxalicum</i>	1
<i>Penicillium pinophilum</i>	3
<i>Penicillium polonicum</i>	6
<i>Penicillium raistrickii</i>	1
<i>Penicillium rolfsii</i>	1
<i>Penicillium sizovae</i>	2
<i>Penicillium spinulosum</i>	2
<i>Penicillium terrigenum</i>	1
<i>Penicillium variabile</i>	2
<i>Phaeosphaeria avenaria</i>	1
<i>Phanerochaete chrysosporium</i>	2
<i>Pithomyces chartarum</i>	3
<i>Pleospora papaveracea</i>	1
<i>Scedosporium boydii</i>	1
<i>Pseudallescheria ellipsoidea</i>	1
<i>Rhizomucor pusillus</i>	2
<i>Rhizomucor variabilis</i>	2
<i>Trichoderma reesei</i>	2
<i>Tyromyces fissilis</i>	3

database to comply with the recommendations we have issued in a previous study [6].

Identification of the strains from the assessment panel

MALDI-TOF MS-based identification

Liquid culture technology used in the Bruker protocol being not available in the two centers that participated to the study, culturing was performed on solid agar. After culturing the molds on solid Sabouraud Chloramphenicol Gentamicin agar plates (pH 5.6; glucose 40 g/L) (OXOID, Dardilly, France) for at least 48 h at 30 °C and the dermatophytes on solid Sabouraud Actidione (Bio Rad) at 27 °C, the strains were treated as described by Cassagne et al.[2], and as advised by the manufacturer : the colonies were gently scrapped using a scalpel blade, and the fungal material (approximately 2–3 mm in diameter) was suspended in a microtube containing 900 µL of anhydrous ethyl alcohol (Carlo Erba SDS, Val de Reuil, France) and

300 µL of sterile water (Water HPLC, Prolabo, BDH, Fontenay-sous-Bois, France). After a 10-min centrifugation step at 13,000 rpm, the pellet was resuspended in 10 µL of formic acid (Sigma-Aldrich, Lyon, France). After 5 min of incubation, 10 µL of acetonitrile (Prolabo BDH) was added. The suspension was then centrifuged at 13,000 rpm for 2 min, and four spots of 1 µL of protein extract per isolate were deposited. Last, the samples were covered with 1 µL of HCCA (α -cyano-4-hydroxycinnamic acid) matrix (Sigma-Aldrich, Lyon, France). MS acquisition was performed either in Marseille or Brussels using a Microflex LT system. All raw spectra were then centralized in Marseille and compared to the two reference databases. For all 1688 tested spectra (four spots for each strain) and each database, three possible scores were collected: the highest score corresponding to a true MS identification result at the species level if available, the highest score corresponding to a true identification at the

genus level but false at the species level if available, and in all cases, the highest score false at the genus level.

DNA sequence-based identification

All strains belonging to the assessment panel were identified via DNA sequencing, which is the current gold standard in fungal identification. Nucleotide sequence analysis was performed either in Marseille ($n = 224$) or Brussels ($n = 198$). The rRNA ITS2 region was sequenced for each strain (using primer sequences ITS3 – GCA TCG ATG AAG AAC GCA GC and ITS4c – TCC TCC GCT TAT TGA TAT GC). An additional locus was analyzed to identify particular taxa: the partial beta-tubulin gene (primer sequences: Bt2A – GGT AAC CAA ATC GGT GCT GCT TTC and Bt2B – ACC CTC AGT GTA GTG ACC CTT GGC) was sequenced for *Aspergillus*, *Penicillium* and *Scedosporium* species, while the elongation factor (primer sequences: EF1 – ATG GGT AAG GAR GAC AAG AC and EF2 – GGA RGT ACC AGT SAT CAT GTT) was sequenced for *Fusarium* species. The hybridization temperatures applied for these primers were 54 °C for the ITS and beta-tubulin genes and 58 °C for the elongation factor gene. The DNA sequence-based identification criteria were as follows: a sequence longer than 350 base pairs and at least 99% homology with the NCBI and CBS nucleotide databases [35, 36].

Identification algorithms

The following 10 distinct algorithms were used to identify the 422 isolates included in the panel.

A: Only one spot (i.e., the first of the four spots) is taken into account.

B: The first two spots are taken into account.

– B₁: Only the identification corresponding to the higher of the two scores is taken into account, plotted and categorized.

– B₂: To be accepted at the species level, both identifications corresponding to the two spots must be identical, and the same rule applies for the identification at the genus level.

C: The first three spots are taken into account.

– C₁: Only the identification corresponding to the highest of the three scores is taken into account, plotted and categorized.

– C₂: To be categorized as a concordant identification, the identification corresponding to the highest score must be identical to at least one of the other two identifications (at either the species or genus level).

– C₃: To be categorized as a concordant identification, the three identifications corresponding to the three spots must be identical (at either the species or genus level).

D: All four spots are taken into account.

– D₁: Only the identification corresponding to the highest of the four scores is taken into account, plotted and categorized.

– D₂: To be categorized as a concordant identification, the identification corresponding to the highest score must be identical to at least one of the other three identifications (at either the species or genus level).

– D₃: To be categorized as a concordant identification, the identification corresponding to the highest score must be identical to at least two

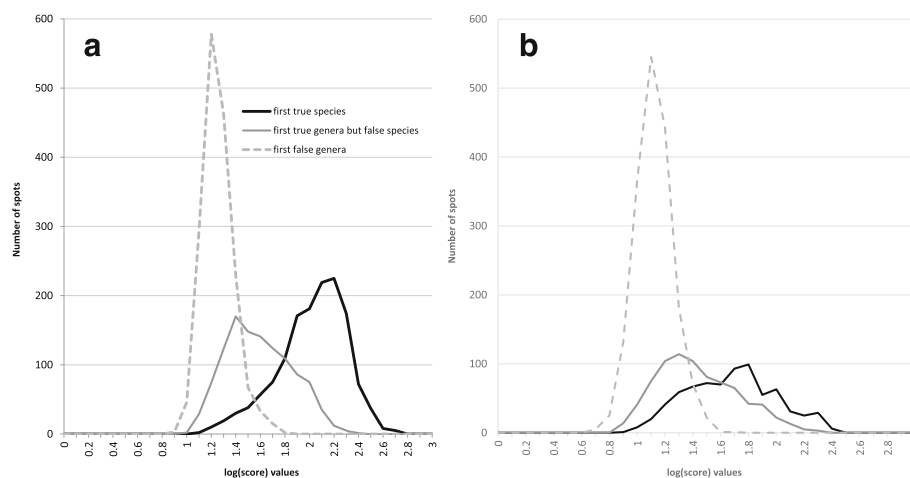


Fig. 1 Highest log(score) value distribution of the 1688 tested spectra. **a** considering the first accurate species identification (dark line), the first accurate genus identification (gray line), and the first false genus identification (dotted line) using the in-house reference database. **b** considering the first accurate species identification (dark line), the first accurate genus identification (gray line), and the first false genus identification (dotted line) using the Bruker reference database

of the other three identifications (at either the species or genus level).

- D₄: To be categorized as a concordant identification, the four identifications corresponding to the four spots must be identical (at either the species or genus level).

Based on the individual spectrum results, 10 identification algorithms (or combinations of parameters) were tested per strain. For each spectrum, only the first identification (either true or false at the species and genus level) and the corresponding scores were collected. Each of the 10 algorithms was then tested using the identification thresholds 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2 and 2.3 for the entire panel consisting of 422 strains. The same process was applied for a reduced panel of 280 isolates that were represented at least three times in the in-

house reference database and for another reduced panel of 191 isolates that were represented at least three times in the commercial database by Bruker.

Statistical analyses

An initial identification assessment was carried out by categorizing the first identification for each spot into three classes: (1) concordant at the species level with the DNA-based identification, (2) concordant at the genus level but not at the species level with the DNA-based identification, or (3) discordant at the genus level with the DNA-based identification. We then determined the best rules for accepting identification at the genus and species levels. For each threshold, algorithm, panel and database, we calculated the percentage of submitted strains that fulfilled the criteria for identification as well as

Table 3 Comparison of the 10 different algorithms tested on the 422 strains included in the assessment panel against the Bruker reference database

Threshold		A	B1	B2	C1	C2	C3	D1	D2	D3	D4
LS 1.5	%accepted	43.71	50.36	38.95	47.74	47.03	31.83	51.31	51.31	44.66	30.40
	species PPV	0.46	0.47	0.43	0.51	0.51	0.46	0.50	0.50	0.48	0.42
	genus PPV	0.65	0.67	0.57	0.70	0.70	0.61	0.69	0.69	0.66	0.54
LS 1.6	%accepted	35.15	40.62	29.22	40.38	39.67	25.18	44.42	44.42	38.00	24.23
	species PPV	0.57	0.58	0.57	0.61	0.60	0.58	0.57	0.57	0.56	0.53
	genus PPV	0.80	0.83	0.76	0.83	0.83	0.77	0.79	0.79	0.78	0.68
LS 1.7	%accepted	28.27	33.73	22.33	33.49	32.78	19.48	35.15	35.15	29.45	16.39
	species PPV	0.71	0.70	0.74	0.73	0.72	0.76	0.72	0.72	0.73	0.78
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 1.8	%accepted	17.58	23.75	19.48	24.47	23.99	17.58	25.89	25.89	22.33	15.20
	species PPV	0.77	0.74	0.76	0.73	0.72	0.74	0.74	0.74	0.74	0.78
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 1.9	%accepted	11.88	16.63	14.25	16.39	15.91	12.59	17.58	17.58	15.68	12.11
	species PPV	0.78	0.77	0.77	0.77	0.76	0.75	0.77	0.77	0.74	0.78
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.0	%accepted	7.84	10.21	9.74	10.45	10.21	9.03	10.45	10.45	9.98	8.79
	species PPV	0.79	0.79	0.78	0.82	0.81	0.79	0.80	0.80	0.79	0.81
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.1	%accepted	4.51	5.46	4.99	5.94	5.94	5.46	5.70	5.70	5.70	5.23
	species PPV	0.79	0.78	0.76	0.80	0.80	0.78	0.75	0.75	0.75	0.77
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.2	%accepted	3.33	4.04	3.56	3.56	3.56	3.09	3.80	3.80	3.80	3.56
	species PPV	0.86	0.88	0.87	0.87	0.87	0.85	0.81	0.81	0.81	0.80
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.3	%accepted	2.14	2.38	1.90	1.19	1.19	0.95	1.19	1.19	1.19	0.95
	species PPV	0.78	0.80	0.75	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

% accepted: percentage of submitted strains that fulfilled the criteria for identification at either the genus or species level, depending on the applied threshold; species PPV: positive predictive value at the species level; genus PPV: positive predictive value at the genus level

the positive predictive value (PPV) of the identification result at the genus and species levels.

Results

MALDI-TOF MS-based identification

Using our in-house database to identify the 422 strains included in the panel, a clear separation was observed between the highest scores corresponding to true identifications at the species level and the highest scores corresponding to a false genus identification. Conversely, it was markedly more difficult to define a satisfying cut-off that differentiated between the highest scores corresponding to an accurate species identification and the highest scores corresponding to an accurate genus but false species identification (Fig. 1a). Using the commercial database led to a similar observation,

although the separation between the three curves was less pronounced (Fig. 1b).

Algorithm performance and identification score threshold

Analyses of the identification performances for the 422-strain panel with the Bruker database and the in-house database. The percentage of identified strains and PPVs at the species and genus levels obtained using the 10 distinct identification algorithms were assessed on the 422-strain panel using the Bruker database (Table 3) as well as the in-house reference database (Table 4).

Notable differences in successful identification rates were highlighted between the 10 algorithms (up to 22% difference between the best identification rate and the worst one for one defined threshold). In particular, with both references databases, the algorithm that selected

Table 4 Comparison of the 10 different algorithms tested on the 422 strains included in the assessment panel against the in-house reference database

Threshold		A	B1	B2	C1	C2	C3	D1	D2	D3	D4
LS 1.5	%accepted	90.97	93.35	81.00	92.64	91.69	73.63	94.54	94.30	88.36	72.45
	species PPV	0.80	0.81	0.82	0.83	0.83	0.85	0.82	0.82	0.82	0.83
	genus PPV	0.90	0.92	0.90	0.93	0.93	0.93	0.92	0.92	0.92	0.91
LS 1.6	%accepted	87.41	90.02	77.67	89.55	88.60	70.78	90.26	90.02	84.32	68.65
	species PPV	0.84	0.84	0.86	0.85	0.86	0.89	0.86	0.86	0.86	0.88
	genus PPV	0.94	0.95	0.94	0.96	0.96	0.96	0.97	0.97	0.97	0.96
LS 1.7	%accepted	81.95	85.27	73.16	85.99	85.27	68.17	87.41	87.17	81.47	65.80
	species PPV	0.89	0.89	0.91	0.89	0.89	0.92	0.89	0.89	0.89	0.92
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 1.8	%accepted	76.96	81.00	71.97	79.33	78.86	65.56	81.71	81.47	77.20	64.61
	species PPV	0.89	0.89	0.91	0.90	0.91	0.92	0.90	0.90	0.90	0.92
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 1.9	%accepted	68.88	74.11	67.46	70.78	70.31	61.28	74.11	73.87	70.55	60.57
	species PPV	0.90	0.89	0.90	0.90	0.91	0.92	0.89	0.89	0.89	0.91
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.0	%accepted	56.77	63.18	57.72	63.18	62.71	55.58	63.66	63.42	61.28	55.11
	species PPV	0.92	0.91	0.92	0.91	0.92	0.92	0.91	0.91	0.91	0.91
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.1	%accepted	43.47	50.36	46.08	49.64	49.41	43.94	50.59	50.36	48.93	44.18
	species PPV	0.94	0.92	0.93	0.92	0.93	0.93	0.92	0.92	0.92	0.92
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.2	%accepted	29.69	35.15	32.30	33.02	32.78	29.93	34.44	34.20	33.97	31.12
	species PPV	0.95	0.93	0.93	0.92	0.93	0.94	0.92	0.92	0.92	0.92
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.3	%accepted	14.25	18.29	17.34	16.86	16.86	16.15	18.05	18.05	18.05	17.34
	species PPV	0.95	0.95	0.95	0.93	0.93	0.93	0.92	0.92	0.92	0.92
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

% accepted: percentage of submitted strains that fulfilled the criteria for identification at either the genus or species level, depending on the applied threshold; species PPV: positive predictive value at the species level; genus PPV: positive predictive value at the genus level

the best results among four spots (D1) increased the accepted identification rate (i.e., percentage of assessed strains that fulfilled the criteria for identification at either the genus or species level) ($A < B1 < C1 < D1$) without notably changing the identification PPV. By contrast, the inter-spot concordance criteria did not increase the identification PPV at the species level and reduced the accepted result rate ($D4 < D3 < D2 \approx D1$). The main difference between the results obtained with the two databases was the percentage of accepted identifications, which was markedly lower with the Bruker database than with the in-house reference database (up to 59.8% difference between the two databases). Additionally, while the PPV for genus level identification was 100% for both databases using the 1.7 threshold, regardless of the other decision criteria applied, the PPV for species level identification was better with our in-house database than with the

Bruker database (from 0.13 units for the D4 algorithm to 0.18 units for the A1 and B1 algorithms). Using the in-house reference database, the D1 combination with a threshold at 1.7 yielded 87.41% of accepted identifications, with a species identification PPV of 0.89.

We noted an increase in the species PPV with threshold values higher than 1.7, but the loss in percentage of identification is important. Using the identification score threshold of 2.00, as recommended by the manufacturer, led to a 23.75% decrease in accepted identifications and only a 2% increase in the PPV. The “highly probable species identification” criterion by Bruker (i.e., a score above the threshold of 2.3) yielded a species PPV of 0.92 but only 18.05% of accepted identifications (76 strains) using our in-house database. The identification results were even less acceptable when using the commercial database,

Table 5 List of the 39 identification failures obtained using the Bruker database and the D1 algorithm with a score above the 1.7 threshold

Identification via molecular biology	Nb of occurrences	Identification D1	Score D1	Number of references in the Bruker reference database
<i>Aspergillus tubingensis</i>	10	<i>Aspergillus niger</i>	2.329/2.161/2.102/1.928/2.238/1.955/ 2.333/1.983/1.987/1.829	0
<i>Fusarium proliferatum</i>	2	<i>Fusarium verticillioides</i>	1.845/1.718	5
<i>Curvularia hawaiiensis</i>	2	<i>Curvularia pallescens</i>	1.879/2.079	0
<i>Penicillium brasilianum</i>	2	<i>Penicillium discolor</i>	1.92/1.887	0
<i>Arthrimum arundinis</i>	1	<i>Arthrimum phaeospermum</i>	1.883	0
<i>Aspergillus quadrilineatus</i>	1	<i>Aspergillus nidulans</i>	2.065	0
<i>Aspergillus calidoustus</i>	4	<i>Aspergillus ustus</i>	1.775/2.076/1.771/1.731	0
<i>Fusarium sacchari</i>	1	<i>Fusarium proliferatum</i>	1.742	0
<i>Aspergillus subolivaceus</i>	1	<i>Aspergillus flavus</i>	1.985	0
<i>Scopulariopsis brevicaulis</i>	2	<i>Scopulariopsis brumptii</i>	1.781/1.925	9
<i>Fusarium verticillioides</i>	1	<i>Fusarium proliferatum</i>	1.953	2
<i>Curvularia lunata</i>	1	<i>Curvularia pallescens</i>	2.039	1
<i>Aspergillus creber</i>	1	<i>Aspergillus versicolor</i>	2.103	0
<i>Aspergillus sydowii</i>	2	<i>Aspergillus versicolor</i>	1.884/1.897	1
<i>Aspergillus chevalieri</i>	2	<i>Aspergillus hollandicus</i>	1.84/1.766	0
<i>Fusarium equiseti</i>	1	<i>Fusarium incarnatum</i>	1.701	1
<i>Trichophyton rubrum</i>	1	<i>Trichophyton mentagrophytes</i>	1.701	6
<i>Aspergillus alabamensis</i>	1	<i>Aspergillus terreus</i>	1.739	0
<i>Aspergillus fischeri</i>	1	<i>Aspergillus fumigatus</i>	1.898	0
<i>Trichoderma orientale</i>	1	<i>Trichoderma longibrachiatum</i>	1.802	0
<i>Trichoderma reesei</i>	1	<i>Trichoderma koningii</i>	1.865	0

with only 44 strains (10.45%) reaching the 2.0 threshold and 5 strains (1.19%) reaching the 2.3 threshold.

Description of the misidentifications obtained with both databases

For 32 strains, the identification best score using the D1 algorithm was higher than the 1.7 threshold, although the species level result was incorrect using the in-house database. For all of these strains, proper identification at the species level was impeded by the close similarity of the MALDI-TOF MS spectra of some species belonging to the same species complex (as it is the case with *Fusarium solani* and its closely related species *Fusarium petrolophilum*). The details of these strains are presented in Table 5. The same finding was

observed for 39 misidentifications obtained using the Bruker database (Table 6).

Analyses of the identification performances for the restricted panels with the Bruker database and the in-house database

The robustness of the identification algorithms was tested by performing the same analysis on a restricted panel including only the strains that were represented by at least three distinct strains in the reference spectrum libraries (the restricted panel used with the Bruker database ($n = 191$) is different than the one used with the in-house database ($n = 280$) due to the differences in databases composition). As expected, the results were better than those obtained with the 422-strain panel (Tables 7 and 8), with a higher percentage of results reaching the thresholds (from

Table 6 List of the 32 identification failures obtained using the in-house database and the D1 algorithm with a score above the 1.7 threshold

Identification via molecular biology	Nb of occurrences	Identification D1	Score D1	Number of references in the in-house reference database
<i>Acremonium strictum</i>	3	<i>Acremonium sclerotigenum</i>	1.952/2.042/2.126	1
<i>Alternaria alternata</i>	1	<i>Alternaria soliaegyptiaca</i>	2.211	1
<i>Aspergillus sclerotiorum</i>	1	<i>Aspergillus persii</i>	1.828	2
<i>Cladosporium halotolerans</i>	1	<i>Cladosporium sphaerospermum</i>	1.975	0
<i>Cladosporium pseudocladosporioides</i>	1	<i>Cladosporium cladosporioides</i>	1.969	1
<i>Aspergillus nidulans</i>	1	<i>Aspergillus echinulatus</i>	2.203	5
<i>Aspergillus quadrilineatus</i>	1	<i>Aspergillus nidulans</i>	2.339	2
<i>Aspergillus hollandicus</i>	1	<i>Aspergillus heterocaryoticus</i>	2.496	1
<i>Aspergillus rubrobrunneus</i>	1	<i>Aspergillus hollandicus</i>	2.325	1
<i>Fusarium solani</i>	2	<i>Fusarium petrolophilum</i>	1.981/2.081	7
<i>Fusarium solani</i>	1	<i>Acremonium falciforme</i>	2.036	7
<i>Trichoderma orientale</i>	1	<i>Trichoderma longibrachiatum</i>	2.026	0
<i>Mucor irregularis</i>	1	<i>Rhizomucor variabilis</i>	2.235	0
<i>Penicillium asturianum</i>	1	<i>Penicillium oxalicum</i>	1.953	1
<i>Penicillium cecidicola</i>	1	<i>Penicillium funiculosum</i>	2.067	1
<i>Penicillium chrysogenum</i>	1	<i>Penicillium nalgiovense</i>	2.328	6
<i>Penicillium glabrum</i>	1	<i>Penicillium purpurescens</i>	2.249	4
<i>Penicillium nalgiovense</i>	1	<i>Penicillium chrysogenum</i>	2.17	3
<i>Talaromyces purpurogenus</i>	1	<i>Talaromyces amestolkiae</i>	2.39	1
<i>Penicillium rubrum</i>	1	<i>Talaromyces amestolkiae</i>	2.069	2
<i>Pleospora papaveracea</i>	1	<i>Ulocladium oudemansii</i>	2.115	0
<i>Pseudallescheria boydii</i>	1	<i>Scedosporium apiospermum</i>	1.721	3
<i>Scedosporium apiospermum</i>	2	<i>Scedosporium boydii</i>	1.746/2.249	5
<i>Scopulariopsis acremonium</i>	1	<i>Scopulariopsis brevicaulis</i>	2.178	1
<i>Talaromyces ramulosus</i>	1	<i>Penicillium cecidicola</i>	2.375	0
<i>Talaromyces stollii</i>	1	<i>Penicillium funiculosum</i>	2.13	2
<i>Trichoderma reesei</i>	1	<i>Trichoderma citrinoviride</i>	1.727	0
<i>Trichophyton rubrum</i>	1	<i>Trichophyton violaceum</i>	1.95	12

Table 7 Performance comparison of the 10 different algorithms tested on the sub-panel of 191 strains that are represented by three or more strains in the Bruker reference library

Threshold		A1	B1	B2	C1	C2	C3	D1	D2	D3	D4
LS 1.5	%accepted	56.02	61.26	48.17	58.12	57.07	38.74	63.35	63.35	53.93	35.08%
	species PPV	0.63	0.66	0.59	0.71	0.71	0.64	0.69	0.69	0.66	0.58
	genus PPV	0.69	0.72	0.64	0.75	0.74	0.68	0.74	0.74	0.72	0.60
LS 1.6	%accepted	46.07	49.74	36.65	49.74	48.69	31.41	56.54	56.54	47.64	28.80
	species PPV	0.76	0.81	0.77	0.83	0.83	0.78	0.77	0.77	0.75	0.71
	genus PPV	0.84	0.88	0.84	0.87	0.87	0.83	0.82	0.82	0.81	0.73
LS 1.7	%accepted	38.74	43.98	30.89	43.46	42.41	26.18	46.60	46.60	38.74	20.94
	species PPV	0.91	0.92	0.92	0.95	0.95	0.94	0.93	0.93	0.92	0.98
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 1.8	%accepted	24.61	31.94	27.23	31.94	31.41	23.04	33.51	33.51	28.80	18.32
	species PPV	0.98	0.95	0.94	0.97	0.97	0.95	0.98	0.98	0.98	1.00
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 1.9	%accepted	17.28	23.04	19.90	21.47	20.94	15.71	23.56	23.56	20.42	15.71
	species PPV	0.97	0.95	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.0	%accepted	12.04	15.71	14.66	14.14	14.14	11.52	13.61	13.61	13.09	11.52
	species PPV	1.00	0.97	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.1	%accepted	6.81	7.85	6.81	8.38	8.38	7.33	7.33	7.33	7.33	6.81
	species PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.2	%accepted	5.24	6.81	5.76	5.76	5.76	4.71	5.24	5.24	5.24	4.71
	species PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.3	%accepted	3.14	3.66	2.62	2.09	2.09	1.57	2.09	2.09	2.09	1.57
	species PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

% accepted: percentage of submitted strains that fulfilled the criteria for identification at either the genus or species level depending on the applied threshold; species PPV: positive predictive value at the species level; genus PPV: positive predictive value at the genus level

3.30% increase in percentage of accepted identification with B1 to 24.94% increase with D4 when selecting the 1.7 threshold) and a reduced number of misidentifications. For example, using the in-house database, the species PPV for the D1 algorithm with a 1.7 score threshold on the comprehensive 422-strain panel increased from 0.89 to 0.97 when testing a sub-panel including only the 280 strains for which species were represented by at least three distinct strains. Similar trends were observed when using the commercial database (increase in the species PPV from 0.72 to 0.93).

Twenty-two species (180 isolates) of this study are represented by at least three references in both databases. We performed a per species comparison of the efficiency of the two databases for those 22 species, using the D1 algorithm. Results are compiled in Table 9. For the total of the 180 isolates, the percentage of accepted identifications is lower with the Bruker

database than with the in-house database (47.22% vs 91.67%). The comparison of the species PPV obtained with both databases highlighted weaknesses for four species (*A. alternata*, *F. solani*, *P. chrysogenum* and *S. apiospermum*) with the in-house database (due to misidentifications with closely related species), and for five species (*A. parasiticus*, *F. proliferatum*, *S. brevicaulis*, *T. interdigitale* and *T. rubrum*) with the Bruker database. The species PPV was equal to 1.00 and identical between the two databases for 13 species of this group.

Discussion

As shown in this study and several others, MALDI-TOF MS has recently been optimized to identify filamentous fungi at the species level, provided that an appropriate database is available [2, 4, 6–8]. The assessment panel used in this study comprised isolates belonging to 126

Table 8 Comparison of the 10 different algorithms tested on a sub-panel of 280 strains that are represented by three or more strains in the in-house reference library

Threshold		A	B1	B2	C1	C2	C3	D1	D2	D3	D4
LS 1.5	%accepted	93.57	96.07	95.36	97.14	97.14	95.36	98.21	98.21	97.14	95.71
	species PPV	0.93	0.94	0.92	0.94	0.94	0.91	0.95	0.93	0.92	0.90
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 1.6	%accepted	90.00	92.50	92.50	93.93	93.93	93.21	95.36	95.36	94.64	93.93
	species PPV	0.94	0.94	0.92	0.95	0.95	0.91	0.96	0.95	0.93	0.91
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 1.7	%accepted	86.07	88.57	88.57	91.07	91.07	90.36	92.14	92.14	91.43	90.71
	species PPV	0.95	0.95	0.92	0.96	0.95	0.92	0.97	0.95	0.94	0.91
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 1.8	%accepted	81.07	82.86	82.86	83.93	83.93	83.21	86.07	86.07	85.71	85.36
	species PPV	0.96	0.96	0.93	0.97	0.97	0.93	0.97	0.96	0.95	0.92
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 1.9	%accepted	69.64	75.36	75.36	76.79	76.79	76.07	79.29	79.29	78.93	78.57
	species PPV	0.97	0.96	0.95	0.97	0.97	0.93	0.97	0.96	0.95	0.93
	genus PPV	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.0	%accepted	60.00	65.00	65.00	68.57	68.57	67.86	70.36	70.36	70.00	69.64
	species PPV	0.98	0.97	0.95	0.97	0.97	0.94	0.97	0.96	0.95	0.94
	genus PPV	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.1	%accepted	46.79	54.64	54.64	57.14	57.14	56.79	59.64	59.64	59.29	59.29
	species PPV	0.98	0.97	0.95	0.98	0.98	0.94	0.98	0.97	0.96	0.94
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.2	%accepted	30.36	36.79	36.79	40.00	40.00	39.64	42.50	42.50	42.14	42.14
	species PPV	0.99	0.97	0.95	0.97	0.97	0.95	0.98	0.97	0.97	0.96
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS 2.3	%accepted	15.71	19.29	19.29	21.43	21.43	21.43	22.50	22.50	22.50	22.50
	species PPV	0.97	0.98	0.96	0.97	0.97	0.95	0.97	0.95	0.95	0.95
	genus PPV	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

% accepted: percentage of submitted strains that fulfilled the criteria for identification at either the genus or species level depending on the applied threshold; species PPV: positive predictive value at the species level; genus PPV: positive predictive value at the genus level

different species, including many difficult-to-identify isolates. Consequently, the identification rate was low compared with previously published studies, especially studies utilizing the commercial database. Using such a challenging panel, it was possible to define a threshold discriminating correct identifications at the species level from false identifications at the genus level, although identifying a threshold that differentiated between a correct and false identification at the species level among closely related species was not feasible. Our findings indicate that the best identification results (PPV at the species level and percentage of accepted identifications) were obtained when applying a decisional algorithm in which only the highest score of four spots was taken into account, and the identification was accepted if the score was at least 1.7. This threshold is notably lower than that recommended for bacteria identification,

and has already been advised by several authors for yeast identification [32, 33]. Furthermore, we also observed a marked decrease in percentage of identification between a threshold of 1.7 and 2.0. When the threshold was set too high, we noted a 20% loss in identification efficiency and only a small benefit of 2% in the differentiation of certain close species. Even if the Bruker commercial database yielded a markedly lower percentage of accepted identifications (probably due to the fact that the database was built using liquid culture media, that is does not frequently involve more than two or three references per species, and that there are errors of reference labeling), the PPV remained correct for the spectra reaching the 1.7 threshold. This result corroborates the findings of several authors who have suggested that lowering the identification score thresholds would improve MALDI-TOF MS-based fungal identification efficiency

Table 9 Comparison of the identification efficiency of the two databases per species, using the D1 algorithm of identification with a 1.7 LS threshold, for 180 strains that are represented by three or more strains in both reference databases

	nb of isolates	In-House database		Bruker database	
		% accepted	species PPV	% accepted	species PPV
<i>Absidia corymbifera</i>	2	100.00	1.00	50.00	1.00
<i>Alternaria alternata</i>	2	100.00	0.50	100.00	1.00
<i>Aspergillus flavus</i>	15	100.00	1.00	73.33	1.00
<i>Aspergillus fumigatus</i>	20	95.00	1.00	45.00	1.00
<i>Aspergillus nidulans</i>	8	100.00	1.00	75.00	1.00
<i>Aspergillus niger</i>	4	100.00	1.00	50.00	1.00
<i>Aspergillus parasiticus</i>	2	100.00	1.00	0.00	Not Relevant
<i>Aspergillus terreus</i>	10	100.00	1.00	30.00	1.00
<i>Aspergillus versicolor</i>	3	66.67	1.00	33.33	1.00
<i>Exophiala dermatitidis</i>	2	50.00	1.00	50.00	1.00
<i>Fusarium oxysporum</i>	21	80.95	1.00	19.05	1.00
<i>Fusarium proliferatum</i>	10	100.00	1.00	60.00	0.33
<i>Fusarium solani</i>	7	57.14	0.25	28.57	1.00
<i>Microsporium canis</i>	2	100.00	1.00	50.00	1.00
<i>Purpureocillium lilacinus</i>	3	66.67	1.00	66.67	1.00
<i>Paecilomyces variotii</i>	3	66.67	1.00	66.67	1.00
<i>Penicillium chrysogenum</i>	27	100.00	0.96	33.33	1.00
<i>Scedosporium apiospermum</i>	11	100.00	0.82	81.82	1.00
<i>Schizophyllum commune</i>	4	100.00	1.00	100.00	1.00
<i>Scopulariopsis brevicaulis</i>	15	86.67	1.00	46.67	0.71
<i>Trichophyton interdigitale</i>	1	100.00	1.00	0.00	Not Relevant
<i>Trichophyton rubrum</i>	8	87.50	0.86	37.50	0.67
Total Isolates	180	91.67	0.95	47.22	0.92

Nb of isolates : number of isolates from the panel of strain that belong to each species. % accepted: percentage of submitted strains that fulfilled the criteria for identification at either the genus or species level; species PPV: positive predictive value at the species level

[7, 17, 32, 33]. However, we strongly advise against a threshold below 1.7, which could yield an increase in false identifications, in particular at the genus level, as shown in Tables 3 and 4.

The major problem encountered with the data set obtained from the assessment panel was the resemblance between the MALDI-TOF MS spectra of some species. For 32 strains, the similarity in MALDI-TOF MS spectra yielded an incorrect identification at the species level, although these identification errors systematically corresponded to taxonomically close species often belonging to the same species complex (except for *Pleospora papaveracea*, which was identified as *Ulocladium oudemansi*). Among these 32 strains (27 species), only a few species are known to be pathogenic to humans, of which none are major human pathogens. Regarding the pathogenic species that were misidentified in our study, the treatment would not vary from that prescribed based on molecular identification, and the differentiation of these species is not clinically relevant.

The remaining species are often considered to be a result of contamination, and their antifungal susceptibility is unknown. Interestingly, raising the identification score threshold to 2.3 did not entirely resolve this identification issue of taxonomically close species, as six strains were incorrectly characterized despite a score above 2.3, even using our extensive in-house database. Thus, in our study, one could question the relevance of capacity and necessity to distinguish between these closely related species as none of the misidentifications implied a change in treatment.

Our results also show that the key step in efficient identification was to increase the number of replicate spots per strain. Testing four spots of each filamentous fungus yielded better identification performance, with an increase of 5% to 7% of accepted strain identifications. Such spot replications did not markedly increase the workload, as it took only a few seconds to perform the additional spots. However, the replicates may increase the per-strain acquisition time in the laboratory workflow. The four-spot

method has been applied in the routine workflow at the Marseille Mycology Laboratory for the past five years, which demonstrates that this procedure is easily applicable in hospital laboratories to identify filamentous fungi isolated from clinical samples and from hospital environment samples. Multiplying the number of deposits per strains imply a small increase in the cost per strain as only one reagent (HCCA matrix) has to be used more extensively. The decisional algorithm implying concordance of the identification results of two to four spots did not enhance identification performance, regardless of the threshold selected. Furthermore, the D2, D3 and D4 algorithms tended to lower the percentage of validated identifications by disregarding identification results that would have been correct using the D1 decisional algorithm.

When focusing on the species that were represented by at least three distinct strains in the reference spectrum library, identification efficacy markedly increased. This result stresses the importance of a comprehensive reference spectrum library for the identification of filamentous fungi, including dermatophytes, as previously reported by our group [6], and as previously demonstrated for yeasts [32]. The multiplication of strains per species as well as subcultures per strain allowed us to cover the variations seen within the environment for filamentous fungi.

Our in-house reference database, which yielded markedly better results compared with the commercial Bruker database, was constructed mainly with collection strains and with some clinical strains isolated in our laboratory and obtained during routine analyses. These clinical strains were later included in the BCCM/IHEM collection in Brussels. Nevertheless, even our enhanced database still requires optimization to identify rare species. Indeed, the poor representation of rare species in the database may explain the identification percentage of 87.4% obtained with this challenging panel of 422 strains (92.1% with the 280 strains represented by more than three strains in the reference database), compared with the extremely good identification results of common species (98.1%) exposed by Gautier et al. [21].

Conclusion

The growing number of publications on MALDI-TOF MS-based fungal identification indicates that this technology is widely used and applicable to the routine laboratory workflow. MALDI-TOF MS-based identification of fungi yields exceedingly more accurate results compared with morphology-based analyses [21]. Furthermore, this technology is less expensive and easier than the current DNA based-identification gold standard, with a turnaround time allowing achieving the analysis more rapidly. Currently, the major limitation of this approach is the lack of a comprehensive and efficient reference spectrum library. We acknowledge that in-

house reference databases that are non-open-source, such as ours, are of little benefit to the global scientific community. Therefore, we plan to deploy an accessible version of our database along with an identification algorithm that can be queried online to assist scientific teams to expand the identification of fungal species. Experiments implying a multicenter approach are ongoing to finalize this project.

Additional file

Additional file 1: Table S1. The species included in the in-house reference database (as labeled by the BCCM/IHEM or Marseille Mycology Laboratory staff at the time of database development). Nb str: number of strains per genus; nb sp: number of species belonging to each genus. The number in brackets represents the number of strains corresponding to each species. (DOCX 52 kb)

Acknowledgements

We especially thank Sandy Moore for proofreading the English version of the manuscript.

Funding

None.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

ACN and PB collected the data. ACN performed the analyses. ACN and RP interpreted the data. ACN, RP, PB, SR and MH were major contributors in writing the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Author details

¹Laboratoire de Parasitologie-Mycologie, CHU Timone, Université de la Méditerranée, Marseille, France. ²Service of Mycology and Aerobiology, BCCM/IHEM Fungal Collection, Scientific Institute of Public Health, Brussels, Belgium.

Received: 10 November 2016 Accepted: 17 January 2017

Published online: 31 January 2017

References

1. Marinach-Patrice C, Lethuillier A, Marly A, Brossas J-Y, Gené J, Symoens F, Datry A, Guarro J, Mazier D, Hennequin C. Use of mass spectrometry to identify clinical *Fusarium* isolates. *Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis.* 2009;15:634–42.
2. Cassagne C, Ranque S, Normand A-C, Fourquet P, Thiebault S, Planard C, Hendrickx M, Piarroux R. Mould routine identification in the clinical laboratory by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *PLoS One.* 2011;6:e28425.
3. De Carolis E, Posteraro B, Lass-Flörl C, Vella A, Florio AR, Torelli R, Girmenia C, Colozza C, Tortorano AM, Sanguinetti M, Fadda G. Species identification of *Aspergillus*, *Fusarium* and *Mucorales* with direct surface analysis by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis.* 2012;18:475–84.

4. Lau AF, Drake SK, Calhoun LB, Henderson CM, Zelazny AM. Development of a clinically comprehensive database and a simple procedure for identification of molds from solid media by matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *J Clin Microbiol*. 2013;51:828–34.
5. L'Ollivier C, Cassagne C, Normand A-C, Bouchara J-P, Contet-Audonneau N, Hendrickx M, Fourquet P, Coulibaly O, Piarroux R, Ranque S. A MALDI-TOF MS procedure for clinical dermatophyte species identification in the routine laboratory. *Med Mycol*. 2013;51:713–20.
6. Normand A-C, Cassagne C, Ranque S, L'Ollivier C, Fourquet P, Roesems S, Hendrickx M, Piarroux R. Assessment of various parameters to improve MALDI-TOF MS reference spectra libraries constructed for the routine identification of filamentous fungi. *BMC Microbiol*. 2013;13:76.
7. Ranque S, Normand A-C, Cassagne C, Murat J-B, Bourgeois N, Dalle F, Gari-Toussaint M, Fourquet P, Hendrickx M, Piarroux R. MALDI-TOF mass spectrometry identification of filamentous fungi in the clinical laboratory. *Mycoses*. 2014;57:135–40.
8. Schulthess B, Ledermann R, Mouttet F, Zbinden A, Bloemberg GV, Böttger EC, Hombach M. Use of the Bruker MALDI Biotyper for identification of molds in the clinical mycology laboratory. *J Clin Microbiol*. 2014;52:2797–803.
9. Li TY, Liu BH, Chen YC. Characterization of *Aspergillus* spores by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom RCM*. 2000;14:2393–400.
10. Welham KJ, Domin MA, Johnson K, Jones L, Ashton DS. Characterization of fungal spores by laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom RCM*. 2000;14:307–10.
11. Chen H-Y, Chen Y-C. Characterization of intact *Penicillium* spores by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun Mass Spectrom RCM*. 2005;19:3564–8.
12. Dong H, Kempfner J, Marchetti-Deschmann M, Kubicek CP, Allmaier G. Development of a MALDI two-layer volume sample preparation technique for analysis of colored conidia spores of *Fusarium* by MALDI linear TOF mass spectrometry. *Anal Bioanal Chem*. 2009;395:1373–83.
13. Kempfner J, Marchetti-Deschmann M, Mach R, Druzhinina IS, Kubicek CP, Allmaier G. Evaluation of matrix-assisted laser desorption/ionization (MALDI) preparation techniques for surface characterization of intact *Fusarium* spores by MALDI linear time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom RCM*. 2009;23:877–84.
14. Tao J, Zhang G, Jiang Z, Cheng Y, Feng J, Chen Z. Detection of pathogenic *Verticillium* spp. using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom RCM*. 2009;23:3647–54.
15. DE Respíns S, Vogel G, Benagli C, Tonolla M, Petrini O, Samuels G. MALDI-TOF MS of *Trichoderma*: A model system for the identification of microfungi. *Mycol Prog*. 2010;9:79–100.
16. Coulibaly O, Marinach-Patrice C, Cassagne C, Piarroux R, Mazier D, Ranque S. *Pseudallescheria/Scedosporium* complex species identification by Matrix-Assisted Laser Desorption Ionization Time-Of-Flight Mass Spectrometry. *Med Mycol Off Publ Int Soc Hum Anim Mycol*. 2011;49:621–6.
17. Del Chierico F, Masotti A, Onori M, Fiscarelli E, Mancinelli L, Ricciotti G, Alghisi F, Dimiziani L, Manetti C, Urbani A, Muraca M, Putignani L. MALDI-TOF MS proteomic phenotyping of filamentous and other fungi from clinical origin. *J Proteomics*. 2012;75(11):3314–30.
18. Nenoff P, Erhard M, Simon JC, Muylowa GK, Herrmann J, Rataj W, Gräser Y. MALDI-TOF mass spectrometry - a rapid method for the identification of dermatophyte species. *Med Mycol*. 2013;51:17–24.
19. Packeu A, Hendrickx M, Beguin H, Martiny D, Vandenberg O, Detandt M. Identification of the *Trichophyton mentagrophytes* complex species using MALDI-TOF mass spectrometry. *Med Mycol*. 2013;51:580–5.
20. Packeu A, De Bel A, L'Ollivier C, Ranque S, Detandt M, Hendrickx M. Fast and accurate identification of dermatophytes by matrix-assisted laser desorption/ionization-time of flight mass spectrometry: validation in the clinical laboratory. *J Clin Microbiol*. 2014;52:3440–3.
21. Gautier M, Ranque S, Normand A-C, Becker P, Packeu A, Cassagne C, L'Ollivier C, Hendrickx M, Piarroux R. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry: revolutionizing clinical laboratory diagnosis of mould infections. *Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis*. 2014;20:1366–71.
22. Erhard M, Hipler U-C, Burmester A, Brakhage AA, Wöstemeyer J. Identification of dermatophyte species causing onychomycosis and tinea pedis by MALDI-TOF mass spectrometry. *Exp Dermatol*. 2008;17:356–61.
23. de Respíns S, Tonolla M, Pranghofer S, Petrini L, Petrini O, Bosshard PP. Identification of dermatophytes by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Med Mycol*. 2013;51:514–21.
24. Verwer PEB, van Leeuwen WB, Girard V, Monnin V, van Belkum A, Staab JF, Verbrugh HA, Bakker-Woudenberg la JM, van de Sande WWJ. Discrimination of *Aspergillus lentulus* from *Aspergillus fumigatus* by Raman spectroscopy and MALDI-TOF MS. *Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol*. 2014;33:245–51.
25. De Respíns S, Monnin V, Girard V, Welker M, Arsac M, Cellière B, Durand G, Bosshard PP, Farina C, Passera M, Van Belkum A, Petrini O, Tonolla M. Matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry using the Vitek MS system for rapid and accurate identification of dermatophytes on solid cultures. *J Clin Microbiol*. 2014;52:4286–92.
26. Lévesque S, Dufresne PJ, Soualhine H, Domingo M-C, Bekal S, Lefebvre B, Tremblay C. A Side by Side Comparison of Bruker Biotyper and VITEK MS: Utility of MALDI-TOF MS Technology for Microorganism Identification in a Public Health Reference Laboratory. *PLoS One*. 2015;10:e0144878.
27. Iriart X, Lavergne R-A, Fillaux J, Valentin A, Magnaval J-F, Berry A, Cassaing S. Routine identification of medical fungi by the new Vitek MS matrix-assisted laser desorption/ionization-time of flight system with a new time-effective strategy. *J Clin Microbiol*. 2012;50:2107–10.
28. McMullen AR, Wallace MA, Pincus DH, Wilkey K, Burnham CA. Evaluation of the Vitek MS Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry System for Identification of Clinically Relevant Filamentous Fungi. *J Clin Microbiol*. 2016;54:2068–73.
29. Alshawa K, Beretti J-L, Lacroix C, Feuillade M, Dauphin B, Quesne G, Hassouni N, Nassif X, Bougnoux M-E. Successful identification of clinical dermatophyte and *Neoscytalidium* species by matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *J Clin Microbiol*. 2012;50:2277–81.
30. Alanio A, Beretti J-L, Dauphin B, Mellado E, Quesne G, Lacroix C, Amara A, Berche P, Nassif X, Bougnoux M-E. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for fast and accurate identification of clinically relevant *Aspergillus* species. *Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis*. 2011;17:750–5.
31. Theel ES, Hall L, Mandrekar J, Wengenack NL. Dermatophyte identification using matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *J Clin Microbiol*. 2011;49:4067–71.
32. Vlek A, Kolecka A, Khayhan K, Theelen B, Groenewald M, Boel E, Multicenter Study Group, Boekhout T. Interlaboratory comparison of sample preparation methods, database expansions, and cutoff values for identification of yeasts by matrix-assisted laser desorption/ionization-time of flight mass spectrometry using a yeast test panel. *J Clin Microbiol*. 2014;52:3023–9.
33. Van Herendael BH, Bruynseels P, Bensaïd M, Boekhout T, De Baere T, Surmont I, Mertens AH. Validation of a modified algorithm for the identification of yeast isolates using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). *Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol*. 2012;31:841–8.
34. Becker PT, de Bel A, Martiny D, Ranque S, Piarroux R, Cassagne C, Detandt M, Hendrickx M. Identification of filamentous fungi isolates by MALDI-TOF mass spectrometry: clinical evaluation of an extended reference spectra library. *Med Mycol*. 2014;52:826–34.
35. Irinyi L, Serena C, Garcia-Hermoso D, Arabatzis M, Desnos-Ollivier M, Vu D, Cardinali G, Arthur I, Normand A-C, Giraldo A, da Cunha KC, Sandoval-Denis M, Hendrickx M, Nishikaku AS, de Azevedo Melo AS, Merseguet KB, Khan A, Parente Rocha JA, Sampaio P, da Silva Briones MR, e Ferreira RC, de Medeiros Muniz M, Castañón-Olivares LR, Estrada-Barcenas D, Cassagne C, Mary C, Duan SY, Kong F, Sun AY, Zeng X, et al. International Society of Human and Animal Mycology (ISHAM)-ITS reference DNA barcoding database—the quality controlled standard tool for routine identification of human and animal pathogenic fungi. *Med Mycol*. 2015;53:313–37.
36. Balajee SA, Borman AM, Brandt ME, Cano J, Cuenca-Estrella M, Dannaoui E, Guarro J, Haase G, Kibbler CC, Meyer W, O'Donnell K, Petti CA, Rodriguez-Tudela JL, Sutton D, Velegriki A, Wickes BL. Sequence-Based Identification of *Aspergillus*, *Fusarium*, and *Mucorales* Species in the Clinical Mycology Laboratory: Where Are We and Where Should We Go from Here? *J Clin Microbiol*. 2009;47:877–84.