Validation of MALDI-TOF MS Biotyper database optimized for anaerobic bacteria: The ENRIA project

A. C. M. Veloo, H. Jean-Pierre, U. S. Justesen, T. Morris, E. Urban, I. Wybo, M. Kostrzewa, A. W. Friedrich, H. Shah, E. Nagy

Anaerobe 54 (2018) 224–230

Contents lists available at ScienceDirect

Anaerobe

journal homepage: www.elsevier.com/locate/anaerobe

Clinical microbiology

Validation of MALDI-TOF MS Biotyper database optimized for anaerobic bacteria: The ENRIA project

A.C.M. Veloo [a, *], H. Jean-Pierre [b, c], U.S. Justesen [d], T. Morris [e], E. Urban [f], I. Wybo [g], M. Kostrzewa [h], A.W. Friedrich [a], on behalf of the ENRIA workgroup

[a] University of Groningen, University Medical Center Groningen, Department of Medical Microbiology, Groningen, The Netherlands
[b] Centre Hospitalier Universitaire de Montpellier, Hôpital Arnaud de Villeneuve, Laboratoire de Bactériologie, 371 Avenue du Doyen Gaston Giraud, 34295 Montpellier Cedex 5, France
[c] Université Montpellier 1, UMR5119 ECOSYM, Equipe Pathogènes Hydriques Santé Environnements, UMR 5569 Hydrosciences, UFR Pharmacie, 15 Avenue Charles Flahault, 34093 Montpellier Cedex 5, France
[d] Department of Clinical Microbiology, Odense University Hospital, Odense, Denmark
[e] UK Anaerobe Reference Unit, Public Health Wales Microbiology, Cardiff, UK
[f] Institute of Clinical Microbiology, University of Szeged, Hungary
[g] Department of Microbiology and Infection Control, Universitair Ziekenhuis Brussel, Brussels, Belgium
[h] Bruker Daltonics, Bremen, Germany

ARTICLE INFO

Article history:
Received 30 November 2017
Received in revised form
6 March 2018
Accepted 9 March 2018
Available online 12 March 2018

Handling Editor: Elisabeth Nagy

Keywords:
ENRIA project
Anaerobic bacteria
Optimization
MALDI-TOF MS
Validation

ABSTRACT

Within the ENRIA project, several 'expertise laboratories' collaborated in order to optimize the identification of clinical anaerobic isolates by using a widely available platform, the Biotyper Matrix Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS). Main Spectral Profiles (MSPs) of well characterized anaerobic strains were added to one of the latest updates of the Biotyper database db6903; (V6 database) for common use. MSPs of anaerobic strains nominated for addition to the Biotyper database are included in this validation. In this study, we validated the optimized database (db5989 [V5 database] + ENRIA MSPs) using 6309 anaerobic isolates.

Using the V5 database 71.1% of the isolates could be identified with high confidence, 16.9% with low confidence and 12.0% could not be identified. Including the MSPs added to the V6 database and all MSPs created within the ENRIA project, the amount of strains identified with high confidence increased to 74.8% and 79.2%, respectively. Strains that could not be identified using MALDI-TOF MS decreased to 10.4% and 7.3%, respectively. The observed increase in high confidence identifications differed per genus. For *Bilophila wadsworthia*, *Prevotella* spp., gram-positive anaerobic cocci and other less commonly encountered species more strains were identified with higher confidence. A subset of the non-identified strains (42.1%) were identified using 16S rDNA gene sequencing. The obtained identities demonstrated that strains could not be identified either due to the generation of spectra of insufficient quality or due to the fact that no MSP of the encountered species was present in the database. Undoubtedly, the ENRIA project has successfully increased the number of anaerobic isolates that can be identified with high confidence. We therefore recommend further expansion of the database to include less frequently isolated species as this would also allow us to gain valuable insight into the clinical relevance of these less common anaerobic bacteria.

1. Introduction

The introduction of Matrix Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) has had a great impact on the identification of anaerobic bacteria isolated from human clinical specimens, providing more timely and accurate identification compared with phenotypic methods [3,4,16]. Several studies have been performed on the performance of MALDI-TOF MS for the identification of anaerobic bacteria. The general conclusion is that the databases of the MALDI-TOF MS systems need

* Corresponding author.
E-mail address: a.c.m.veloo@med.umcg.nl (A.C.M. Veloo).

https://doi.org/10.1016/j.anaerobe.2018.03.007

optimization for the identification of anaerobic bacteria [4,9,17]. To accomplish this we formed the European Network for the Rapid Identification of Anaerobes (ENRIA). Within this project, seven European laboratories, with expertise in anaerobic bacteriology, collaborated in order to collect clinical isolates for addition to the MALDI-TOF MS database. As a quality requirement, the ENRIA group set itself the goal of obtaining at least five main spectral profiles (MSPs) for each species [16].

The ENRIA group felt that it was vital for the project to benefit the microbiology community worldwide and therefore we setup a collaboration with Bruker Daltonics, Bremen, Germany. At the time of going to press the first batch of MSPs of ENRIA strains have been added to the db6903 database (V6 database) of the Bruker MALDI-TOF MS Biotyper system. In this study we present the validation of this optimized database for the identification of anaerobic bacteria. Genera/species that require further optimization are also highlighted.

2. Material and methods

2.1. Bacterial strains

Over a period of six months all anaerobic bacteria isolated from human clinical specimens were identified using the db5989 database (V5 database) and the V5 database plus ENRIA MSPs, created and supplied by Bruker Daltonics. The latter consisted of two parts: the confirmed ENRIA MSPs (the ones included in the last update of the Biotyper V6 database) and all ENRIA MSPs (the confirmed ENRIA MSPs and the MSPs nominated for adding to the Biotyper database). An overview of these MSPs is shown in the supplemental data, Table 1. This was performed by all 'expertise laboratories': University Medical Center Groningen (UMCG), Groningen, The Netherlands; Center Hospitalier Universitaire de Montpellier, Montpellier, France; Odense University Hospital, Odense, Denmark; UK Anaerobe Reference Unit (UKARU), Public Health Wales Microbiology, Cardiff, United Kingdom; University of Szeged, Szeged, Hungary and University Hospital Brussels, Brussels, Belgium.

2.2. Identification

The MALDI-TOF MS Biotyper (Bruker Daltonics, Bremen, Germany) was utilized at each laboratory to perform the measurements as described previously [18]. Briefly, bacterial cells in the log phase of their growth were spotted on a stainless target twice using a toothpick. One spot was overlaid with 1 µl HCCA matrix (α-cyano-4-hydroxycinnamic acid in 50% acetonitrile/2.5% trifluoro-acetic acid) and left to dry at ambient temperature. An on target extraction was performed on the other spot by first overlaying the spot with 1 µl 70% formic acid. Immediately after drying at ambient temperature, the spot was overlaid with 1 µl HCCA matrix. Each laboratory performed the measurements as part of their daily routine, using the standard settings. Obtained log scores were interpreted as advised by the manufacturer. A log score of $\geq$2 was considered an identification with high confidence, a log score of $\geq$1.7 and $<$2 as an identification with low confidence and a log score of $<$1.7 as no reliable identification.

2.3. Data interpretation

The obtained identifications were divided into 3 groups. No reliable identification (or no identification), identification with low confidence and an identification with a high confidence. The latter two were interpreted as a reliable genus identification and a reliable species identification, respectively. For several species an identification to the subspecies level was given by the Biotyper, for example for *Fusobacterium nucleatum* and *Actinomyces neuii*. However, for the data interpretation only the identification to the species level was considered. Furthermore, we are aware of the fact that several species of *Veillonella* are difficult to separate from each

Table 1
The distribution of the different genera isolated by all expertise laboratories, used for the validation of the optimized database.

Genus	No. of strains	% of total
Bacteroides spp.	934	14.8%
Cutibacterium spp.	647	10.3%
Prevotella spp.	582	9.2%
Clostridiodes spp.	413	6.5%
Finegoldia magna[a]	412	6.5%
Peptoniphilus spp.	349	5.5%
Actinomyces spp.	306	4.9%
Fusobacterium spp.	303	4.8%
Parvimonas micra	244	3.9%
Veillonella spp.	241	3.8%
Anaerococcus spp.	230	3.6%
Clostridium spp.	225	3.6%
Peptostreptococcus spp.	130	2.1%
Porphyromonas spp.	129	2.0%
Dialister spp.	69	1.1%
Eggerthella lenta[a]	65	1.0%
Atopobium spp.	58	0.9%
Parabacteroides spp.	54	0.8%
Bifidobacterium spp.	52	0.8%
Campylobacter spp.	48	0.7%
Solobacterium moorei[a]	41	0.6%
Slackia spp.	31	0.5%
Propionimicrobium spp.	30	0.5%
Propionibacterium spp.	26	0.4%
Bilophila wadsworthia[a]	24	0.4%
Alloscardovia spp.	16	0.3%
Hungatella spp.	16	0.3%
Helcococcus spp.	15	0.2%
Murdochiella asaccharolytica[a]	13	0.2%
Paeniclostridium spp.	10	0.2%
Flavonifractor plautii[a]	9	0.1%
Lachnoanaerobaculum spp.	9	0.1%
Filifactor spp.	9	0.1%
Alistipes spp.	8	0.1%
Eubacterium spp.	8	0.1%
Acidaminococcus spp.	7	0.1%
Mogibacterium spp.	7	0.1%
'Fenollaria massiliensis'	7	0.1%
Odoribacter spp.	7	0.1%
Olsenella spp.	7	0.1%
Peptococcus niger[a]	7	0.1%
Eggerthia catenaformis[a]	7	0.1%
Desulfovibrio spp.	6	0.1%
Paraclostridium spp.	5	0.1%
Collinsella spp.	4	0.1%
Ruminococcus spp.	4	0.1%
Sutterella spp.	4	0.1%
Bulleidia extructa[a]	3	0.0%
Leptotrichia spp.	3	0.0%
Hathewaya spp.	2	0.0%
Terrisporobacter spp.	2	0.0%
Dielma fastidiosa[a]	2	0.0%
Moryella indoligenes[a]	2	0.0%
Selenomonas spp.	2	0.0%
Intestinibacter spp.	1	0.0%
Alloprevotella spp.	1	0.0%
Butyricimonas spp.	1	0.0%
Cetobacterium spp.	1	0.0%
Megasphaera spp.	1	0.0%
Tissierella spp.	1	0.0%
Blautia spp.	1	0.0%
no identification	458	7.3%
Total	6309	

[a] Genera represented by only one species are represented in the table using the species name.

other by MALDI-TOF MS and 16S rDNA gene sequencing, namely, *Veillonella dispar*, *Veillonella parvula*, *Veillonella denticariosi* and *Veillonella rogosae*. Strains identified as any of these four species are named as *Veillonella* spp., regardless of the log score obtained during the measurement. Also some species of *Bacteroides* cannot be differentiated from each other. Therefore strains identified as *Bacteroides ovatus* or *Bacteroides xylanisolvens, Bacteroides thetaiotaomicron*or or *Bacteroides faecis* and *Bacteroides vulgatus* or *Bacteroides dorei* were identified as either one of the two. The MALDI-TOF MS also has difficulties with differentiation of *Fusobacterium naviforme* from *Fusobacterium nucleatum*. These two strains were identified as *F. nucleatum/naviforme*. Furthermore, we noticed that it is also difficult to differentiate *Porphyromonas asaccharolytica* from *Porphyromonas uenonis*. Strains identified as either one of these species were designated as *P. asaccharolytica/uenonis*. In order to maximize the potential opportunity of gaining insight into the clinical relevance of rare anaerobic species, we did not differentiate between valid species and non-valid species.

2.4. 16S rDNA gene sequencing

A subset of the strains that could not be identified using MALDI-TOF MS, were identified using 16S rDNA gene sequencing (193 of 458 (42.1%)). This was performed at the originating laboratory, using their own primers and methods [7,8,20].

3. Results

3.1. Distribution of bacterial strains

The optimized MALDI-TOF MS database was validated using 6309 anaerobic strains isolated from human clinical specimens. The distribution of the genera is shown in Table 1. The most prominent genus is *Bacteroides*, which represented 14.8% of the total number of strains analyzed. The genera *Cutibacterium* (formerly *Propionibacterium* [15]) and *Prevotella*, represented 10.3% and 9.2% of the total isolates respectively. Other notable genera which were included in sizeable numbers include those that belong to the gram-positive anaerobic cocci (GPAC) (*Finegoldia, Peptoniphilus*, *Parvimonas*, *Anaerococcus*, *Peptostreptococcus*, *Murdochiella* and *Peptococcus*) accounting for 21.9% as a group of all the analyzed anaerobes.

3.2. Validation

All results are per genus presented in Table 2. Detailed results per genus are presented in the “Data in Brief” [19]. Of the 6309 analyzed strains, 4485 (71.1%) strains were identified with high confidence, 1064 (16.9%) with low confidence and 760 (12.0%) could not be identified, using the V5 database. Adding the confirmed ENRIA MSPs increased the amount of strains identified with high confidence to 4718 (74.8%) and decreased the number of strains identified with low confidence and with no identification to 937 (14.9%) and 654 (10.4%), respectively. For 19.1% of the strains a higher log score was obtained. Adding all ENRIA MSPs increased the amount of strains identified with high confidence even further, to 4999 (79.2%), whilst substantially decreasing those that identified with low confidence to 852 (13.5%). The number of strains with no identification decreased to 458 (7.3%). Therefore, a higher confidence of identification was obtained for 35.2% of the strains upon adding all ENRIA MSPs.

The observed increase in genus/species identification differed by genus. *Bilophila wadsworthia* was only represented in the V5 database at the genus level as ‘*Bilophila* spp’. The addition of MSPs of clinical isolates resulted in an increase of the number of strains with a log score $\geq$2 from 2 strains to 17 strains, respectively. The addition of MSPs of two species of *Lachnoanaerobaculum*, namely *Lachnoanaerobaculum orale* and *Lachnoanaerobaculum umeaenense*, ultimately resulted in a high confidence identification of all tested strains, whereas previously only 5 of the 9 strains (55.6%) could be identified. Using the V5 database, 79.2% of the *Prevotella* strains were identified with high confidence. This increased to 85.7% and 91.0% by adding the confirmed ENRIA MSPs and all ENRIA MSPs, respectively, partly due to the addition of MSPs of species that were not represented in the database. 69.2% of the strains belonging to the GPAC genera were identified with high confidence using the V5 database. Adding confirmed ENRIA MSPs and all ENRIA MSPs resulted in a high confidence identification for 76.5% and 86.4% of the strains, respectively. This increase was due in particular to the addition of MSPs of species (valid and non-valid) not yet represented in the database, such as *Peptoniphilus duerdenii*, *Peptoniphilus tyrrelliae*, ‘Peptoniphilus rhinitidis’ and *Peptostreptococcus stomatis*. Increasing the number of MSPs of *Anaerococcus vaginalis* gave a significant increase for strains identified with high confidence, from 12.1% to 85.0% using the V5 database and all ENRIA MSPs, respectively. With the latter database, a higher log score was obtained for 100% of the *A. vaginalis* strains. The amount of *Porphyromonas* strains identified with high confidence also increased mostly due to the addition of *P. asaccharolytica/uenonis* ENRIA MSPs (from 53.6% using the V5 database to 80.0% by including all ENRIA MSPs).

Several less common species could now be identified using the V5 database optimized with all ENRIA MSPs, including strains of *Dielma fastidiosa*, *Moryella indoligenes*, ‘Fenollaria massiliensis’, *Filifactor alocis* and *Odoribacter splanchnicus*.

3.3. 16S rDNA gene sequencing

Of the 6309 analyzed strains, 458 were not identified using MALDI-TOF MS. To assess the reason for this, 193 of these strains were sequenced by the originating laboratories. Species for which no identification was obtained for more than one strain are shown in Table 3. A portion of these strains gave no spectrum during MALDI-TOF MS analyses and these were mostly species belonging to the genus *Actinomyces* and *Propionibacterium/Cutibacterium* (data not shown). For other species no identification was obtained due to the fact that they are not represented in the MALDI-TOF MS database; e.g. *Eisenbergiella tayi*, *Anaerovorax odorimutans* and *Akkermansia muciniphila*. Other species are represented in the database, but not identified by MALDI-TOF MS; e.g. *Dialister pneumosintes*, *Eggerthella lenta*, *Atopobium minutum* and *Anaerococcus prevotii*. Species of which only one strain was encountered and not represented in the database included; *Anaerocolumna aminovalerica*, ‘Casaltella massiliensis’, *Clostridium tetanomorphum,* and *Haloimpatiens lingqiaonensis*.

4. Discussion

In this study, we validated the Bruker MALDI-TOF MS database, optimized for anaerobic bacteria by utilizing a large set of anaerobic strains isolated from human clinical specimens. Several European laboratories specializing in anaerobic bacteria collaborated as part of the validation, allowing us to include strains of species less commonly encountered in human infections from within the collaborators private collections. Identifications with high confidence increased by 3.7%, from 71.1% to 74.8%, using the V5 database and confirmed ENRIA MSPs. Using V5 and all ENRIA MSPs, identifications with high confidence increased by 8.1%, from 71.1% to 79.2%.

The three most prevalent genera were *Bacteroides* (13.5%), *Cutibacterium* (9.4%) and *Prevotella* (8.7%). For the first two genera,

Table 2
The MALDI-TOF MS results of the validation using the optimized Biotyper database for anaerobic bacteria.

Strains (6309)	V5 database[a]			V5 database + ENRIA (confirmed)			higher score	V5 database + ENRIA (all MSPs)		higher score
	≤1.7	1.7–2	≥2	≤1.7	1.7–2	≥2		1.7–2	≥2	
Acidaminococcus spp. (7)		1	6		1	6	5	1	6	5
Alistipes spp. (8)	1		7			8	1		8	5
Alloscardovia spp. (16)		2	14		2	14		2	14	
Atopobium spp. (58)	9	9	40	9	9	40	12	8	50	25
Alloprevotella spp. (1)	1			1				1		1
Bifidobacterium spp. (52)		12	40		12	40	10	12	40	10
Bilophila wadsworthia (24)	7	15	2	2	5	17	20	7	17	22
Bulleidia extructa (3)			3			3			3	
Butyricimonas spp. (1)		1			1			1		
Collinsella spp. (4)			4			4			4	
Campylobacter spp. (48)	4	16	28	2	17	29	5	18	30	7
Cetobacterium spp. (1)	1				1		1	1		1
Desulfovibrio spp. (6)	6			3	1	2	3	2	4	6
Dialister spp. (69)		7	62		4	65	60	4	65	60
Dielma fastidiosa (2)	2			2					2	2
Eubacterium spp. (8)		1	7		1	7		1	7	
Eggerthella lenta (65)		10	55		10	55		10	55	
Eggerthia catenaformis (7)			7			7	3		7	5
Flavonifractor plautii (9)		1	8			9	6		9	6
Helcococcus spp. (15)			15			15	2		15	2
Lachnoanaerobaculum spp. (9)		4	5		4	5			9	5
Leptotrichia spp. (3)[b]			3			3			3	
Megasphaera spp. (1)			1			1			1	
Moryella indoligenes (2)	2				1	1	2	1	1	2
Mogibacterium spp. (7)	7			7				6	1	7
Filifactor spp. (9)	9				1	8	9	1	8	9
'Fenollaria massiliensis' (7)	7			7					7	7
Odoribacter spp. (7)	7			4		3	3		7	7
Olsenella spp. (7)	1	1	5	1		6	5	1	6	6
Ruminococcus spp. (4)		1	3			4	1		4	1
Selenomonas spp. (2)	2					2	2		2	2
Slackia spp. (31)			31			31			31	
Solobacterium moorei (41)		4	37		1	40	32	1	40	32
Sutterella spp. (4)			4			4			4	
Tissierella spp. (1)			1			1			1	
Actinomyces spp. (306)	4	96	206		89	217	59	89	217	78
Veillonella spp. (241)	2	160	79	2	160	79	1	144	97	68
Blautia spp. (1)			1			1			1	
Bacteroides spp. (934)	6	45	883	4	36	894	123	24	910	296
Clostridium spp. (225)		31	194		28	197	20	27	198	33
Paraclostridium spp. (5)		4	1		4	1		4	1	
Clostridioides spp. (413)		17	396		17	396		17	396	
Hungatella spp. (16)			16			16			16	5
Terrisporobacter spp. (2)			2			2			2	1
Paeniclostridium spp. (10)		1	9		1	9		1	9	3
Intestinibacter spp. (1)		1			1			1		
Hathewaya spp. (2)			2			2			2	
Parabacteroides spp. (54)	1	5	48		1	53	32	1	53	33
Prevotella spp. (582)	30	81	471	14	58	510	176	48	534	385
Fusobacterium spp. (303)	7	80	216	7	79	217	22	67	236	108
Anaerococcus spp. (230)	52	92	86	32	61	137	83	39	191	165

(*continued on next page*)

Table 2 (*continued*)

Strains (6309)	V5 database[a]			V5 database + ENRIA (confirmed)			higher score	V5 database + ENRIA (all MSPs)		higher score
	≤1.7	1.7–2	≥2	≤1.7	1.7–2	≥2		1.7–2	≥2	
Finegoldia magna (412)		87	325		87	325		87	325	
Murdochiella asaccharolytica (13)		5	8		4	9	6	4	9	6
Peptoniphilus spp. (349)	66	59	224	32	52	265	85	24	325	298
Peptostreptococcus spp. (130)	31	8	91	31	5	94	73	12	118	105
Peptococcus niger (7)	1	6			2	5	7	2	5	7
Parvimonas micra (244)		20	224		20	224		20	224	
Porphyromonas spp. (129)	36	29	64	36	20	73	49	28	101	89
Cutibacterium spp. (647)		118	529		107	540	287	101	546	304
Propionibacterium spp. (26)		6	20		6	20		6	20	
Propionimicrobium lymphophilum (30)		28	2		28	2		28	2	
no ID (458)	458			458						
Totals (n)	760	1064	4485	654	937	4718	1205	852	4999	2219
%	12,0%	16,9%	71,1%	10,4%	14,9%	74,8%	19,1%	13,5%	79,2%	35,2%

[a] The V5 database is the db5989 of the Biotyper system, which was current at the time of the study.
[b] All three strains were only identified at the genus level with a log score ≥2.

Table 3
An overview of species, of which more than one strain was encountered, which could not be identified using MALDI-TOF MS. The identity of the strain was determined using 16S rDNA gene sequencing.

Species	no. of strains
Actinobaculum massiliense	7
Actinotignum urinale	2
Actinotignum spp.	4
Actinomyces cardiffensis	2
Actinomyces israelii	10
Actinomyces naeslundii	5
Actinomyces odontolyticus	5
Actinomyces oris	2
Actinomyces turicensis	4
Akkermansia muciniphila	2
Anaerococcus spp.	14
Anaerococcus prevotii	2
Anaeroglobus geminatus	2
Anaerovorax odorimutans	2
Atopobium minutum	2
Atopobium spp.	2
'Clostridium neonatale'	2
Clostridium spp.	3
Dialister pneumosintes	2
Eggerthella lenta	2
Eisenbergiella tayi	2
Eubacterium spp.	2
Ezakiella spp.	4
Fastidiophila sanguinis	2
Fusobacterium nucleatum	3
genus nov	2
Jonquetella anthropi	2
Negativicoccus succinivorans	2
Mogibacterium spp.	4
Olsenella spp.	7
Paraeggerthella hongkongensis	2
Peptoniphilus spp.	12
Prevotella buccae	2
Prevotella intermedia/nigrescens	3
Prevotella multiformis	6
Prevotella spp.	7
Cutibacterium acnes	4
Propionibacterium freudenreichii	2
Varibaculum cambriense	6

the number of identifications with high confidence did not increase significantly, as they were already well represented in the MALDI-TOF MS database. Within the V5 database the *Bacteroides* species *B. dorei* and *B. xylanisolvens* are not represented (they are however mentioned as 'cross-matching' in a matching hint of the software), and MSPs of these species are included in the collection of all ENRIA MSPs. As previously described by Pedersen et al. [13], the MALDI-TOF MS is not able to distinguish between *B. vulgatus* and *B. dorei* or between *B. ovatus* and *B. xylanisolvens*. Therefore these species are referred to as *B. vulgatus/dorei* and *B. ovatus/xylanisolvens*.

Wybo et al. [20] demonstrated that optimizing the Biotyper database for the identification of *Prevotella* species resulted in a high confidence identification of clinical isolates of *Prevotella* of 89.2%. More recently Gürsoy et al. [6] validated the identification of species of oral origin and found that all strains were correctly identified at the low confidence level (log scores 1.7–2) and to a high confidence species level for 88.6%. Similar rates of high confidence identification for *Prevotella* species were observed in this study. After optimization with confirmed ENRIA MSPs the level of high confidence identifications increased from 79.2% to 85.7% and finally to 91.0% when all ENRIA MSPs were added to the V5 database. Veloo et al. [17] demonstrated that the number of GPAC strains identified with high confidence increased from 53.6% to 82.1% when the Biotyper database was optimized for GPAC species

and this was partly due to the addition of MSPs of species that were not yet represented in the database. Similarly in this study a significant increase was observed in the number of *A. vaginalis* strains which could, due to optimization, be identified with high confidence as opposed to not being identified at all. After adding all ENRIA MSPs to the V5 database, 86.4% of the GPAC could be identified with high confidence, a similar proportion to the previous study. No studies have hitherto been performed on the validation of MALDI-TOF MS specifically for the identification of *Porphyromonas* species isolated from human clinical specimens, with the exception of *Porphyromonas gingivalis*, as this is an important periodontal pathogen [14]. During our study, we observed difficulties differentiating *P. asaccharolytica* from *P. uenonis*, using both MALDI-TOF MS and 16S rDNA gene sequencing. This can be explained by the fact that the 16S rDNA sequence difference between these two species is less than 2% [5]. Differentiation of these two species using MALDI-TOF MS requires further investigation that is beyond the scope of this publication.

Certain species remain difficult to identify using MALDI-TOF MS. Of the strains in this category from within our study we observed that 18% of these belonged to the genus *Actinomyces* and that these often did not give an identification as no spectrum was obtained. This could be explained by the dry colony morphology of certain species, which hampers the spotting [18], or due to the thick cell wall of this bacterium. In both cases, it is advised to perform a full extraction to overcome these problems. A significant number of *Veillonella* strains were included in the validation. From the obtained log scores for the species *V. dispar*, *V. parvula*, *V. dentacariosi* and *V. rogosae* we observed that the MALDI-TOF MS has difficulty differentiating between these species. Marchandin et al. [10] described the phenomenon of micro-heterogeneity, which indicates scattered nucleotide differences, between the different copies of the 16S rDNA gene present in one strain of *Veillonella*. This makes it difficult to differentiate *V. dispar* and *V. parvula* from each other using solely 16S rDNA gene sequencing. It is advised that for the accurate speciation of strains from the genus *Veillonella* that three different housekeeping genes: dnaK, rpoB and gltA are used [1]. When MALDI-TOF MS is used for the identification of bacteria, mostly the ribosomal proteins are measured [11]. Since several *Veillonella* species cannot be differentiated from each other using 16S rDNA gene sequencing, it is not surprising therefore that this is also the case when MALDI-TOF MS is used for the identification. Further research into the identification of *Veillonella* using MALDI-TOF MS is necessary.

Historically, the identification of anaerobic bacteria has been overlooked by diagnostic laboratories largely due to the difficulties associated with the phenotypic identification systems designed specifically for anaerobes. Also the perception exists that a specialized workflow is required, which would be more time consuming than for aerobic bacteria. Many new species have been proposed recently of which the phenotypic features are poorly described. Therefore, commercial identification systems where the identification tables are not kept up to date are unlikely to cover these species. In order to successfully monitor the clinical relevance of these often unvalidated species we chose to add these to the ENRIA database. An example is 'Fenollaria massiliensis', originally isolated from an osteoarticular sample [12]. This species is not validated and its clinical relevance is unknown. During our study we encountered seven strains of this species, which would previously have gone unrecognized in clinical laboratories. This example demonstrates that it is also important to have unvalidated species validated, especially if we wish to gain insight into the clinical relevance of such unusual new species. The importance of including less common species within the databases of MALDI-TOF MS systems is also emphasized by a study by Bernard et al. [2]. They encountered eight isolates of *E. tayi*, isolated from blood cultures, concluding that this species is a potential pathogen. Creating MSPs for these strains ensured that this species can be identified using MALDI-TOF MS.

The validation of the MALDI-TOF V5 database expanded using clinical isolates collected within the ENRIA project clearly demonstrated that an increase in the percentage of strains identified was achievable. It should be noted that not all species can be differentiated from each other, especially certain species of the genus *Veillonella*. We recommend the continual expansion of the MALDI-TOF MS database with MSPs of less common species, valid and unvalid, in order to gain insight into their clinical relevance.

Acknowledgment

ENRIA is partly funded by InterregIVa (grant number III-1-02=73), for the identification of neglected infectious disease and within the task of reference laboratory in the Dutch-German border region. No direct funding was received from Bruker Daltonics. However, Bruker Daltonics provided knowledge, and expertise for this study is therefore considered an equal partner in ENRIA.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.anaerobe.2018.03.007.

References

[1] F. Aujoulat, P. Bouvet, E. Jumas-Bilak, H. Jean-Pierre, H. Marchandin, *Veillonella seminalis* sp. nov., a novel anaerobic Gram-stain-negative coccus from human clinical samples, and emended description of the genus *Veillonella*, Int. J. Syst. Evol. Microbiol. 64 (2014) 3526–3531.

[2] K. Bernard, T. Burdz, D. Wiebe, B.M. Balcewich, T. Zimmerman, P. Lagacé-Wiens, L.M.N. Hoang, A.M. Bernier, Characterization of isolates of *Eisenbergiella tayi*, a strictly anaerobic Gram-stain variable bacillus recovered from human clinical material in Canada, Anaerobe 44 (2017) 128–132.

[3] A. Bizzini, G. Greub, Matrix-assisted laser desorption ionization time-of-flight mass spectrometry, a revolution in clinical microbial identification, Clin. Microbiol. Infect. 16 (2010) 1614–1619.

[4] E. Culebras, I. Rodríguez-Avial, C. Betriu, M. Gómez, J.J. Picazo, Rapid identification of clinical isolates of *Bacteroides* species by matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry, Anaerobe 18 (2012) 163–165.

[5] S.M. Finegold, M.L. Vaisanen, M. Rautio, E. Eerola, P. Summanen, D. Molitoris, Y. Song, C. Liu, H. Jousimies-Somer, *Porphyromonas uenonis* sp. nov., a pathogen for humans distinct from *P. asaccharolytica* and *P. endodontalis*, J. Clin. Microbiol. 42 (2004) 5298–5301.

[6] M. Gürsoy, I. Harju, J. Matomäki, A. Bryk, E. Könönen, Performance of MALDI-TOF MS for identification of oral *Prevotella* species, Anaerobe 47 (2017) 89–93.

[7] R.A. Hutson, D.E. Thompson, M.D. Collins, Genetic interrelationships of saccharolytic *Clostridium botulinum* types B, E and F and related clostridia as revealed by small-subunit rRNA gene sequences, FEMS Microbiol. Lett. 108 (1993) 103–110.

[8] U.S. Justesen, M.N. Skov, E. Knudsen, H.M. Holt, P. Søgaard, T. Justesen, 16S rRNA gene sequencing in routine identification of anaerobic bacteria isolated from blood cultures, J. Clin. Microbiol. 48 (2010) 946–948.

[9] B. La Scola, P.E. Fournier, D. Raoult, Burden of emerging anaerobes in the MALDI-TOF and 16S rRNA gene sequencing era, Anaerobe 17 (2011) 106–112.

[10] H. Marchandin, C. Teyssier, M.S. de Buochberg, H. Jean-Pierre, C. Carriere, E. Jumas-Bilak, Intra-chromosomal heterogeneity between the four 16S rRNA gene copies in the genus *Veillonella*: implications for phylogeny and taxonomy, Microbiol 149 (2003) 1493–1501.

[11] P.R. Murray, What is new in clinical microbiology-microbial identification by MALDI-TOF MS mass spectrometry: a paper from the 2011 William Beaumont Hospital Symposium on molecular pathology, J. Mol. Diagn. 14 (2012) 419–423.

[12] I. Pagnier, O. Croce, C. Robert, D. Raoult, B. La Scola, Non-contiguous finished genome sequence and description of *Fenollaria massiliensis* gen. nov., sp. nov., a new genus of anaerobic bacterium, Stand Genomic Sci 15 (2014) 704–717.

[13] R.M. Pedersen, E.S. Marmolin, U.S. Justesen, Species differentiation of *Bacteroides dorei* from *Bacteroides vulgatus* and *Bacteroides ovatus* from *Bacteroides xylanisolvens* – back to basic, Anaerobe 24 (2013) 1–3.

[14] T.E. Rams, J.D. Sautter, A. Getreu, A.J. van Winkelhoff, Phenotypic identification of *Porphyromonas gingivalis* validated with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, Microb. Pathog. 94 (2016)

112–116.
[15] C.F.P. Scholz, M. Kilian, The natural history of cutaneous propionibacteria, and reclassification of selected species within the genus *Propionibacterium* to the proposed novel genera *Acidipropionibacterium* gen. nov., *Cutibacterium* gen. nov. and *Pseudopropionibacterium* gen. nov, Int. J. Syst. Evol. Microbiol. 66 (2016) 4422–4432.
[16] P. Seng, M. Drancourt, F. Gouriet, B. La Scola, P.E. Fournier, J.M. Rolain, D. Raoult, Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry, Clin. Infect. Dis. 49 (2009) 543–551.
[17] A.C. Veloo, E.D. de Vries, H. Jean-Pierre, U.S. Justesen, T. Morris, E. Urban, I. Wybo, A.J. van Winkelhoff, The optimization and validation of the Biotyper MALDI-TOF MS database for the identification of gram-positive anaerobic cocci, Clin. Microbiol. Infect. 22 (2016) 793–798.
[18] A.C.M. Veloo, P.E. Elgersma, A.W. Friedrich, E. Nagy, A.J. van Winkelhoff, The influence of incubation time, sample preparation and exposure to oxygen on the quality of the MALDI-TOF MS spectrum of anaerobic bacteria, Clin. Microbiol. Infect. 20 (2014) 1091–1097.
[19] A.C.M. Veloo, H. Jean-Pierre, U.S. Justesen, T. Morris, E. Urban, I. Wybo, M. Kostrzewa, A.W. Friedrich, on behalf of the ENRIA workgroup, An overview of the data obtained during the validation of an optimized MALDI-TOF MS Biotyper database for the identification of anaerobic bacteria, Anaerobe (2018) (submitted for publication).
[20] I. Wybo, O. Soetens, A. De Bel, F. Echalidi, E. Vancutsem, K. Vandoorslaer, D. Piérard, Species identification of clinical Prevotella isolates by matrix assisted laser desorption ionization-time of flight mass spectrometry, J. Clin. Microbiol. 50 (2012) 1415–1418.