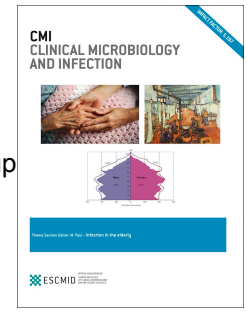
Accepted Manuscript



MALDI-TOF/MS identification of species from the *Acinetobacter baumannii* (Ab) group revisited: inclusion of the novel *A. seifertii* and *A. dijkshoorniae* species

Marta Marí-Almirall, Clara Cosgaya, Paul G. Higgins, Ado Van Assche, Murat Telli, Geert Huys, Bart Lievens, Harald Seifert, Lenie Dijkshoorn, Ignasi Roca, Jordi Vila

PII: S1198-743X(16)30604-8

DOI: 10.1016/j.cmi.2016.11.020

Reference: CMI 788

To appear in: *Clinical Microbiology and Infection*

Received Date: 25 October 2016

Revised Date: 28 November 2016

Accepted Date: 28 November 2016

Please cite this article as: Marí-Almirall M, Cosgaya C, Higgins PG, Van Assche A, Telli M, Huys G, Lievens B, Seifert H, Dijkshoorn L, Roca I, Vila J, MALDI-TOF/MS identification of species from the *Acinetobacter baumannii* (Ab) group revisited: inclusion of the novel *A. seifertii* and *A. dijkshoorniae* species, *Clinical Microbiology and Infection* (2017), doi: 10.1016/j.cmi.2016.11.020.

Original Article

MALDI-TOF/MS identification of species from the *Acinetobacter baumannii* (Ab) group revisited: inclusion of the novel *A. seifertii* and *A. dijkshoorniae* species

Marta Marí-Almirall[1†], Clara Cosgaya[1†], Paul G. Higgins[2,3], Ado Van Assche[4], Murat Telli[5], Geert Huys[6], Bart Lievens[4], Harald Seifert[2,3], Lenie Dijkshoorn[7], Ignasi Roca[1*], and Jordi Vila[1]

[1]Department of Clinical Microbiology and ISGlobal- Barcelona Ctr. Int. Health Res. CRESIB, Hospital Clínic - Universitat de Barcelona, Barcelona, Spain.

[2]Institute for Medical Microbiology, Immunology and Hygiene; University of Cologne, Cologne, Germany

[3]German Centre for Infection Research (DZIF), Partner site Bonn-Cologne, Germany

[4]Laboratory for Process Microbial Ecology and Bioinspirational Management (PME&BIM), Department of Microbial and Molecular Systems (M2S), KU Leuven, Sint-Katelijne-Waver, Belgium

[5]Department of Clinical Microbiology, School of Medicine, Adnan Menderes University, Aydin, Turkey.

[6]BCCM/LMG Bacteria Collection & Laboratory of Microbiology, Faculty of Sciences, Ghent University, Ghent, Belgium.

[7]Department of Infectious Diseases, Leiden University Medical Center, Leiden, The Netherlands

†: These authors equally contributed to this work

***Corresponding author:**

Ignasi Roca. ISGlobal, Barcelona Ctr. Int. Health Res. (CRESIB), Hospital Clínic - Universitat de Barcelona, Barcelona, Spain, CEK building, Spain. Phone: +34932275400; Fax: +34933129410; E-mail: Ignasi.roca@isglobal.org.

Running head: MALDI-TOF/MS identification of *Acinetobacter* spp. revisited

Keywords: *Acinetobacter*, MALDI-TOF/MS, *rpoB*, MLSA, ClinProTools.

ABSTRACT

Objectives: Rapid identification of *Acinetobacter* species is critical since members of the *A. baumannii* (Ab) group differ in antibiotic susceptibility and clinical outcomes. *A. baumannii*, *A. pittii* and *A. nosocomialis* can be identified by MALDI-TOF/MS, while the novel species *A. seifertii* and *A. dijkshoorniae* cannot. Low identification rates for *A. nosocomialis* have also been reported. We evaluated the use of MALDI-TOF/MS to identify isolates of *A. seifertii* and *A. dijkshoorniae* and revisited the identification of *A. nosocomialis* to update the Bruker taxonomy database.

Methods: Species characterisation was performed by *rpoB*-clustering and MLSA. MALDI-TOF/MS spectra were recovered from formic acid/acetonitrile bacterial extracts overlaid with α-cyano-4-hydroxy-cinnamic acid matrix on a MicroflexLT in linear positive mode and 2,000-20,000 *m/z* range mass. Spectra were examined with the ClinProTools v2.2 software. Mean spectra (MSP) were created with the BioTyper software.

Results: Seventy-eight *Acinetobacter* isolates representative of the Ab group were used to calculate the average spectra/species and generate pattern recognition models. Species-specific peaks were identified for all species, and MSPs derived from 3 *A. seifertii*, 2 *A. dijkshoorniae* and 2 *A. nosocomialis* strains were added to the Bruker taxonomy database, allowing successful identification of all isolates using spectra from either bacterial extracts or direct colonies, resulting in a positive predictive value (PPV) of 99.6% (777/780) and 96.8% (302/312), respectively.

Conclusions: The use of post-processing data software identified statistically significant species-specific peaks to generate reference signatures for rapid accurate identification of species within the Ab group, providing relevant information for the clinical management of *Acinetobacter* infections.

INTRODUCTION

The use of matrix-assisted laser desorption ionisation-time of flight mass spectrometry (MALDI-TOF/MS) for the identification of bacterial species has been a major breakthrough in clinical microbiology. MALDI-TOF/MS has proven to be a rapid and accurate methodology highly relevant for the differentiation of closely related bacterial species that are otherwise indistinguishable by conventional phenotypic methods, providing an inexpensive alternative to the laborious and time-consuming molecular identification methods [1].

Former members of the *Acinetobacter baumannii* (Ab) group (*A. baumannii*, *A. nosocomialis* and *A. pittii*) are virtually indistinguishable using conventional phenotypic tests while accurate species differentiation is achieved by sequencing of the RNA polymerase β-subunit (*rpoB*) gene, the DNA gyrase B (*gyrB*) gene and/or by multilocus sequence analysis (MLSA), all of which most likely constitute the current gold standard for molecular identification [2-4].

In a previous work, we evaluated and optimised the use of MALDI-TOF/MS for species identification of the former members of the Ab group and demonstrated that it was an accurate and reliable method [5]. Subsequent MALDI-TOF/MS studies by several other groups together with the recent technological advances in molecular methods (such as whole genome sequencing) have revealed a relative abundance of non-*baumannii* *Acinetobacter* species of the Ab group in clinical specimens, mostly involving *A. nosocomialis* and *A. pittii* isolates [6-11].

In the last few years the taxonomy of the genus *Acinetobacter* has undergone major modifications, with more than 18 new species having been described since 2014 [3]. In particular, two novel pathogenic species, *A. seifertii* and *A. dijkshoorniae*, have recently been included within the Ab group and, like the former members of the group, they can

be best differentiated by molecular methods [3, 12]. Identification of these novel species by MALDI-TOF/MS is not yet possible, since a thorough study that evaluates the distinctness of spectral signatures of all the species within the Ab group and provides reference spectra for the novel species is still lacking. In addition, several studies have shown that while the Bruker MALDI-TOF BioTyper system correctly identifies almost all *A. baumannii* and *A. pittii* isolates, identification rates for *A. nosocomialis* range at about 70%, suggesting that the Bruker database should be updated and further improved to allow efficient identification of all *Acinetobacter* species [8, 13, 14].

The aim of the present study was to perform an in-depth analysis of the spectrum profiles of all the *Acinetobacter* species currently included in the Ab group, and generate reference spectra to allow accurate and reliable identification to the species level by MALDI-TOF/MS.

MATERIALS and METHODS

Bacterial isolates

The present study included 78 isolates belonging to the five *Acinetobacter* species within the Ab group, *A. baumannii* (n=16), *A. nosocomialis* (n=24), *A. pittii* (n=15), *A. dijkshoorniae* (n=12) and *A. seifertii* (n=11), mainly obtained from clinical samples in different geographical locations over a period of 15 years (**Supplementary Table S1**). Isolates were identified at the species level by sequencing of the RNA polymerase β-subunit (*rpoB*) gene and multilocus sequence analysis (MLSA), as described previously [3]. Isolates were preserved at -80°C in 10% skimmed milk until use.

Sample preparation and MALDI-TOF/MS data acquisition

Bacterial cultures were grown overnight on Columbia sheep blood agar (Becton Dickinson, Heidelberg, Germany) at 37ºC and subjected to ethanol-formic acid extraction according to [5].

One microliter of each bacterial extract was spotted onto a MALDI target plate (MSP 96 target ground steel; Bruker Daltonics, Bremen, Germany) and air-dried at room temperature. Each spotted sample was then overlaid with 1 µL of a saturated matrix solution (α-cyano-4-hydroxy-cinnamic acid; Bruker Daltonics) in 50% acetonitrile-2.5% trifluoroacetic acid (Sigma-Aldrich chemical Co., Madrid, Spain) and air-dried.

For MALDI-TOF/MS analysis performed directly from grown bacterial colonies, a small fraction of a single colony was spotted onto the MALDI target plate, carefully spread and subsequently overlaid with 1 µl of matrix.

MALDI-TOF/MS was conducted in a Microflex LT (Bruker Daltonics) benchtop instrument as described previously [5]. Bacterial extracts from all isolates were spotted 5 times onto a MALDI target plate and each spot was measured twice, resulting in 10 mass spectra for each individual isolate. Direct colony samples were spotted twice, and each spot was also measured twice, resulting in 4 mass spectra for each individual isolate.

MALDI-TOF/MS data analysis

Spectra from bacterial extracts were loaded into the ClinProTools software (version 2.2; Bruker Daltonics) and prepared for analysis with the following parameters: 800 resolution, Top Hat baseline subtraction with a 10% minimal baseline width and no data reduction. Null spectra and noise spectra exclusion with a noise threshold of 2.00 were both enabled and spectra grouping was also supported. Peak selection and average peak list calculation ranged from 2,000 to 10,000 mass to charge ratio values (*m/z)*, and

recalibration was performed with a 1,000 parts per million (ppm) maximal peak shift and 30% match to calibrant peaks. Non-recalibrated spectra were excluded.

m/z values from average spectra were identified according to their statistical significance, as determined by the different statistical tests supported by ClinProTools: Anderson-Darling test, t-/ANOVA test and Wilcoxon/Krustal-Wallis test. Informative peaks were those showing a significant difference among all species as described previously [15].

For the generation and validation of pattern recognition models, the 78 isolates were divided into two sets – (i) a reference set containing 40 isolates: *A. baumannii* (n=7), *A. nosocomialis* (n=13), *A. pittii* (n=8), *A. dijkshoorniae* (n=6) and *A. seifertii* (n=6); and (ii) a validation set containing 38 isolates: *A. baumannii* (n=9), *A. nosocomialis* (n=11), *A. pittii* (n=7), *A. dijkshoorniae* (n=6) and *A. seifertii* (n=5). Selection was performed on the grounds of the spectral analysis in order to include as much diversity as possible within both sets, prioritising the reference set whenever an equitable distribution was not possible. Classification models were generated using the genetic algorithm (GA), supervised neural network (SNN), and QuickClassifier (QC) algorithms with default settings. The recognition capability and cross validation values were calculated to demonstrate the reliability and accuracy of the model.

Bacterial identification

Spectra were analysed with the MALDI BioTyper software (version 3.1; Bruker Daltonics) using the pre-processing and BioTyper main spectrum (MSP) identification standard methods (mass range: 2,000 to 20,000 *m/z*) against either the default Bruker database or the association of the Bruker database and our own reference spectra. Accuracy of the identification was determined by a logarithmic score value resulting from the alignment of peaks to the best matching reference spectrum [5].

Bacterial extracts of isolates selected for MSP creation were re-spotted 10 times onto a ground steel target and each spot was measured 3 times. The resulting 30 mass spectra were carefully analysed using the FlexAnalysis software (version 3.4; Bruker Daltonics) to yield a minimum of 20 spectra per isolate with a *m/z* shift of less than 0.05%. Selected spectra were then uploaded onto the MALDI BioTyper to create a single MSP for each isolate with the BioTyper MSP creation standard method.

The MSP dendrogram was constructed using the correlation distance measure with the weighted linkage algorithm settings of the MALDI BioTyper software.

rpoB-based cluster analysis as well as MLSA cluster analysis were performed as described elsewhere [3].

RESULTS

Spectral analysis

Seventy-eight *Acinetobacter* isolates representative of *A. baumannii*, *A. nosocomialis*, *A. pittii*, *A. dijkshoorniae*, and *A. seifertii* were used to identify species-specific biomarker peaks using the Bruker ClinProTools software. Acquired spectra were loaded into ClinProTools and grouped into 5 different classes, one for each *Acinetobacter* species, and the average spectrum for each class was calculated. A detailed spectra analysis of each species was performed in the region between 2,000 and 10,000 *m/z* that concentrated the bulk of mass peaks, and several species-specific peaks ranging from 2876 to 8857 *m/z* values were identified (**Table 1**), as were 6 peaks (4265, 4661, 5175, 6090, 6948 and 9319 *m/z*) that were present in all isolates.

For *A. baumannii*, a biomarker peak unique to this species and present in all isolates was located at 5747 *m/z* (**Figure 1E**), as described previously [5, 14, 16, 17]. Two additional peaks at 4244 and 8485 *m/z*, likely representing the single and double protonation states of the same protein, were also identified in all *A. baumannii* isolates

but these were also shared with some isolates of *A. nosocomialis*, as shown below. Spectra from all *A. baumannii* isolates were correctly identified as *A. baumannii* (100%) by the Bruker BioTyper using the default taxonomy database.

For *A. nosocomialis*, several peaks unique to this species were identified. A pair of peaks at 4069 and 8135 *m/z* (**Figure 1B and 1H**), also likely corresponding to different protonation states of a single protein, were present in all isolates but one, with the latter isolate displaying a shifted version of the pair located at 4084 and 8165 *m/z*, respectively (data not shown), in good agreement with previous reports [14, 17]. A second pair of unique peaks was located at 4180 and 8358 *m/z* but it was present in only 12 out of 24 isolates (**Figure 1C and 1I**).

Bacterial identification using the default Bruker taxonomy database was able to identify as *A. nosocomialis* only 13 out of 24 isolates (54%), while the remaining isolates were misidentified as *A. baumannii* (46%). Similar inconsistent results regarding the identification of *A. nosocomialis* have also been reported by other authors [8, 13, 14]. The *A. nosocomialis* isolates that were correctly classified by the Bruker BioTyper software were clustered together (group I) and compared with those that were misidentified (group II). Spectra within each group showed very similar peak profiles but there were some significant differences between both groups (**Figure 2D**).

As shown in **Figure 2A and 2C**, isolates in *A. nosocomialis* group I and II shared the *A. nosocomialis* species-specific peaks centred around 4069 and 8135 *m/z*. All but one isolate in group I also presented the other two species-specific peaks centred around 4180 and 8358 *m/z*, which were absent among group II profiles. Instead, isolates in group II presented the two peaks at 4244 and 8485 *m/z* that were also present in all *A. baumannii* isolates, as mentioned before.

Despite the clear splitting of all *A. nosocomialis* isolates into two separate groups according to their peak signatures, this clear distinction was not observed by *rpoB*-based clustering (data not shown) or MLSA analysis (**Figure 3**).

For *A. pittii*, the analysis of spectra recognised a higher degree of variability with only a few species-specific peaks shared by a majority of isolates. A major peak at 5777 *m/z* (**Figure 1E**) was identified in all isolates but one, in good agreement with previous reports [14, 16-18], and a second unique peak at 6692 *m/z* was present in 10 out of 15 isolates (**Figure 1F**). Eight out of 15 isolates also showed a pair of peaks located at 4411 and 8821 *m/z*, respectively, again likely representing the differently charged states of a single protein (**Figure 1D and 1J**). These two peaks were also identified in 2 additional isolates although they were shifted to 4347 and 8691 *m/z*, respectively (data not shown). Similarities among *A. pittii* isolates at the spectra level did not correlate with either *rpoB* or MLSA clustering either (data not shown). Species identification using the default Bruker taxonomy database correctly identified all spectra as *A. pittii*, in agreement with previous reports [13, 14].

For *A. dijkshoorniae*, the analysis of spectra identified 4 masses that were present in all isolates and were also unique to this species: 3 major peaks located at 4430, 5788 and 8857 *m/z* values (**Figure 1D, 1E and 1J**), as described previously [3], as well as a smaller peak at 6729 *m/z* (**Figure 1F**). Since there were no reference spectra for this novel *Acinetobacter* species in the Bruker taxonomy database, the best identification matches of spectra from *A. dijkshoorniae* isolates were to *A. pittii* reference spectra. Interestingly, the first two best matches were always to the same *A. pittii* reference spectra (*A. pittii* serovar 18 DSM9341 and serovar 22 DSM9318), with log scores >2.0, while the subsequent best matches against *A. pittii* isolates showed log scores <2.0. It is plausible that the two isolates originally used to create these MSPs belonged to the

novel *A. dijkshoorniae* species. Of note, the MSPs for these isolates were used to screen our *Acinetobacter* collection and led to the identification of 3 isolates that turned out to be *A. dijkshoorniae* by molecular methods.

For *A. seifertii*, a unique peak was present at 7446 *m/z* (**Figure 1G**) in all isolates and several additional species-specific peaks were also identified, albeit with differences depending on the isolates. A pair of peaks located at 4194 and 8385 *m/z* (**Figure 1C and 1I**) was identified in all but 2 isolates that, nevertheless, presented a similar pair but shifted at 4122 and 8244 *m/z* (data not shown). Likewise, another pair of peaks located at 3948 and 7893 *m/z* (**Figure 1A and 1H**) was present in all but 3 isolates, the latter showing a shifted pair at 3985 and 7968 *m/z* (2 isolates) or 3961 and 7920 *m/z* (1 isolate) (data not shown).

As it occurred with *A. dijkshoorniae*, there were no reference spectra for *A. seifertii* in the Bruker taxonomy database either, and the best identification matches were to *A. baumannii* reference spectra. However, the first best match was always to the same *A. baumannii* reference spectra (*A. baumannii* CS_62_1 BRB) with log scores >2.0, while the subsequent best matches showed log scores <2.0. The MSP for *A. baumannii* CS_62_1 BRB was also used to screen our *Acinetobacter* collection and it led to the identification of one isolate that was confirmed as *A. seifertii* by molecular methods, again suggesting that the isolate used to create the MSP *A. baumannii* CS_62_1 BRB most likely belonged to the novel *A. seifertii* species.

Generation and validation of pattern recognition models

Spectra from a reference set of isolates (see Materials and Methods) were uploaded to the ClinProTools software and grouped again into 5 different classes according to each *Acinetobacter* species. The average spectra from each *Acinetobacter* species were used to generate classification models based on the Genetic Algorithm, SNN and Quick

Classifier algorithms to select an optimal set of peaks that allowed correct species allocation of the spectra used for model generation. All three algorithms provided recognition and cross-validation values above 95% and 87%, respectively, suggesting that successful differentiation of all 5 *Acinetobacter* species was possible. Of the three algorithms, the SNN model yielded the highest recognition and cross-validation values (100% and 92.6%, respectively) and was therefore selected to evaluate its ability to classify spectra from isolates not included in the generation of the model (external validation). The SNN model was able to allocate most of the spectra from the 38 isolates of the validation set to their corresponding *Acinetobacter* species, resulting in a positive predictive value (PPV) of 96.8% (**Table 2**).

BioTyper database update and automated identification

As described in Materials and Methods, new BioTyper MSPs were created from representative isolates to account for the intra- and inter-species variability observed. MSPs for *A. seifertii* derived from isolates NIPH 973^{T} (type strain), R00-JV54 and LUH 05789. MSPs for *A. dijkshoorniae* originated from isolates JVAP01^{T} (type strain) and R10-JV222. In addition, we included new MSPs for *A. nosocomialis* that derived from isolates SCOPE 150 and RUH 503, to account for the identification of *A. nosocomialis* isolates belonging to *A. nosocomialis* group II (**Figure 2**).

Cluster analysis of MSPs from all the *Acinetobacter* species within the *Acinetobacter calcoaceticus-Acinetobacter baumannii* complex (which includes the Ab group) grouped MSPs from each *Acinetobacter* species into separate monophyletic clusters (**Figure 4**). Interestingly, the two MSPs from representative isolates of *A. nosocomialis* group II were grouped more closely to *A. baumannii* MSPs than to those of *A. nosocomialis* group I, while still forming a separate clade, also in good agreement with results from the spectral analysis.

Spectra from all 78 isolates were then analysed against a custom database that included the MSPs from all the *Acinetobacter* species within the default Bruker taxonomy database plus the novel reference signatures for *A. seifertii*, *A. dijkshoorniae* and *A. nosocomialis.* As shown in **Table 2**, the allocation of spectra obtained from bacterial extracts to their corresponding *Acinetobacter* species provided sensitivity and specificity values ranging from 98.8-100% and 99.6-100%, respectively, resulting in a PPV of 99.6%. In addition, strains RUH 204 (*A. junii*), RUH 44 (*A. haemolyticus*), RUH 45 (*A. lwoffii*), RUH 3517 (*A. radioresistens*), and RUH 584 (*A. calcoaceticus*), representing a set of reference *Acinetobacter* strains belonging to *Acinetobacter* species other than those included within the Ab group [5], were also correctly identified (data not shown). These results showed the absence of cross-identification between the novel MSPs and other *Acinetobacter* spp.

Likewise, the identification of spectra from direct colonies instead of bacterial extracts yielded sensitivity and specificity values ranging from 91.7-100% and 98.0-100%, respectively, with a PPV of 96.8% (**Table 2**).

DISCUSSION

In the present study we have compared for the first time the spectral profiles of the current members of the Ab group, including the novel *A. seifertii* and *A. dijkshoorniae* species. Spectral analysis has allowed the identification of a conserved set of peaks that are present in all isolates and, therefore, are linked at least to the Ab group. Four of these peaks correspond to 4 out of the 5 peaks described by Sousa *et al.* as being specific to the *Acinetobacter* genus (4662, 5176, 6949 and 9323 *m/z*) [17]. We have found, however, that the peak described by Sousa *et al.* at 7435 *m/z* is present in all species except *A. seifertii*, which instead presents a unique peak at 7446 *m/z*.

The thorough analysis of all spectra has also led to the identification of several peaks that are unique to each *Acinetobacter* species and might serve for identification purposes. The majority of such peaks corroborate previous findings but, nevertheless, remarkable differences have also been found. For instance, previous studies have reported a peak at 2875 *m/z* as specific to *A. baumannii* [5, 19], although Sousa *et al.* reported such peak in *A. nosocomialis* [17]. In the present study, we have identified a peak at 2876 *m/z* in all *A. baumannii* isolates that overlaps with a small intensity peak centred around 2869 *m/z* present in some *A. nosocomialis* isolates. The presence of such a peak might be misleading for identification purposes (data not shown). Likewise, Hsueh *et al.* identified a peak at 2889 *m/z* that was unique to *A. pittii* [14] and in our study this peak is indeed present in all *A. pittii* isolates but it is also identified in several isolates of *A. nosocomialis* and *A. dijkshoorniae*; and a peak at 9542 *m/z* considered as unique to *A. seifertii* by Sousa *et al.* [17] is clearly present in several *A. pittii* and *A. dijkshoorniae* isolates in our study.

In addition, the comprehensive examination of the *A. nosocomialis* isolates has led to their differentiation into two groups according to their spectra profiles. Isolates included in group I contain 4 *A. nosocomialis*-specific peaks while isolates in group II only show two of these peaks but share two additional peaks with *A. baumannii*. Interestingly, of the 5 reference spectra (MSPs) for *A. nosocomialis* included in the Bruker taxonomy database (**Figure 3**), 4 originated from isolates belonging to group I and only one was representative of group II. These differences together with the underrepresentation of group II MSPs in the default BioTyper software might account for the low rates of successful identification of *A. nosocomialis* isolates reported by several authors [8, 13, 14].

So far only two studies have confronted the ambiguous identification of certain isolates from the Ab group, either using alternative sample preparation protocols [16] or by coupling MALDI-TOF/MS with chemometric methods [17]. These novel approaches have certainly improved the differentiation of the former members of the Ab group, but they have failed to provide automated spectra acquisition linked to automated species identification and, therefore, cannot be successfully implemented in routine clinical laboratories. In addition, none of these studies have thoroughly evaluated the identification of the novel members of the Ab group, *A. dijkshoorniae* and *A. seifertii*.

The results from the spectral analysis and the validation of the SNN pattern recognition model in our study suggest that conventional and automated MALDI-TOF/MS identification of all the current members of the Ab group is possible with an updated reference taxonomy database. We have created novel reference signatures (MSPs) to improve the identification of group II *A. nosocomialis* isolates as well as of the novel species *A. dijkshoorniae* and *A. seifertii*. Cluster analysis of the novel MSPs together with those already present in the default Bruker taxonomy database also support the unambiguous identification of all species using this technology. Of note, the MSPs from isolates *A. pittii* serovar 18 DSM9341 and *A. pittii* serovar 22 DSM9318 are clustered together with those of *A. dijkshoorniae*, and the MSP from *A. baumannii* CS_62_1 BRB is clustered together with the MSPs from *A. seifertii* isolates (**Figure 4**), emphasising that the species identification of these isolates should be revisited since the characterisation of novel *Acinetobacter* species.

Bacterial identification by MALDI-TOF/MS using our custom taxonomy database has shown correct identification of all *Acinetobacter* species within the Ab group with sensitivity and specificity values well above 98% when using spectra from bacterial extracts, and above 91% and 98%, respectively, when using spectra directly from

bacterial colonies. It should be noted, however, that the quality of spectra with the use of direct colonies highly depends on technician expertise during sample loading, with identification rates varying greatly.

Inclusion of these novel MSPs into the Bruker taxonomy database should allow rapid automated identification of all the *Acinetobacter* species within the group, contributing to the assessment of the clinical and epidemiological relevance of the different species in the Ab group and, eventually, improving the treatment and management of *Acinetobacter* infections [20].

We acknowledge that the small number of isolates included might have been a limitation in our study, in particular for *A. pittii*. *A. pittii* isolates show the largest variability, both regarding spectra profiles and genetic sequences, and although there is no cross-identification between *A. pittii* and other *Acinetobacter* spp., the inclusion of additional strains may contribute to further delineate this species.

It is also clear from this study that achieving correct identification of bacterial species by MALDI-TOF/MS strongly relies on the accuracy and robustness of the reference database, which needs to be constantly refined and validated on a par with an evolving taxonomic classification.

ACKNOWLEDGEMENTS

We are truly thankful to all the donors shown in **Supplementary Table S1** as well as the Italian APSI-group for kindly providing some of the isolates included in the study.

TRANSPARENCY DECLARATION

This study was supported by the Ministerio de Economía y Competitividad, Instituto de Salud Carlos III, cofinanced by European Regional Development Fund (ERDF) “A Way to Achieve Europe,” the Spanish Network for Research in Infectious Diseases (REIPI RD12/0015), and the Spanish Ministry of Health (grant number FIS 14/0755).

This study was also supported by grant 2014 SGR 0653 from the Departament d'Universitats, Recerca i Societat de la Informació, of the Generalitat de Catalunya, by grant 10030 of the Adnan Menderes University Scientific Research Foundation and by funding from the European Community (MagicBullet, contract HEALTH-F3-2011-278232). C. Cosgaya and M. Marí-Almirall were supported with grants FPU 13/02564 and FPU 14/06357, respectively, from the Spanish Ministry of Education, Culture and Sports. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

These data were partially presented at the 10th International Symposium on the Biology of *Acinetobacter*, Athens, Greece, 2015 (communication no. P 5a).

The authors declare no conflicts of interest.

386 **REFERENCES**

387 [1] Cheng K, Chui H, Domish L, Hernandez D, Wang G. Recent development of
388 mass spectrometry and proteomics applications in identification and typing of bacteria.
389 Proteomics Clin Appl 2016; 10: 346-57.

390 [2] La Scola B, Gundi VA, Khamis A, Raoult D. Sequencing of the *rpoB* gene and
391 flanking spacers for molecular identification of *Acinetobacter* species. J Clin Microbiol
392 2006; 44: 827-32.

393 [3] Cosgaya C, Mari-Almirall M, Van Assche A, Fernandez-Orth D, Mosqueda N,
394 Telli M, *et al.* *Acinetobacter dijkshoorniae* sp. nov., a new member of the *Acinetobacter*
395 *calcoaceticus-Acinetobacter baumannii* complex mainly recovered from clinical
396 samples in different countries. Int J Syst Evol Microbiol 2016; 66: 4105-11.

397 [4] Higgins PG, Lehmann M, Wisplinghoff H, Seifert H. *gyrB* multiplex PCR to
398 differentiate between *Acinetobacter calcoaceticus* and *Acinetobacter* genomic species 3.
399 J Clin Microbiol 2010; 48: 4592-94.

400 [5] Espinal P, Seifert H, Dijkshoorn L, Vila J, Roca I. Rapid and accurate
401 identification of genomic species from the *Acinetobacter baumannii* (Ab) group by
402 MALDI-TOF MS. Clin Microbiol Infect 2012; 18: 1097-103.

403 [6] Wang X, Chen T, Yu R, Lu X, Zong Z. *Acinetobacter pittii* and *Acinetobacter*
404 *nosocomialis* among clinical isolates of the *Acinetobacter calcoaceticus-baumannii*
405 complex in Sichuan, China. Diagn Microbiol Infect Dis 2013; 76: 392-95.

406 [7] Teixeira AB, Martins AF, Barin J, Hermes DM, Pitt CP, Barth AL. First report
407 of carbapenem-resistant *Acinetobacter nosocomialis* isolates harboring IS*Aba1-*
408 *bla*OXA-23 genes in Latin America. J Clin Microbiol 2013; 51: 2739-41.

409 [8] Kishii K, Kikuchi K, Matsuda N, Yoshida A, Okuzumi K, Uetera Y, *et al.*
410 Evaluation of matrix-assisted laser desorption ionization-time of flight mass

411 spectrometry for species identification of *Acinetobacter* strains isolated from blood
412 cultures. Clin Microbiol Infect 2014; 20: 424-30.
413 [9] Zander E, Fernandez-Gonzalez A, Schleicher X, Dammhayn C, Kamolvit W,
414 Seifert H, *et al.* Worldwide dissemination of acquired carbapenem-hydrolysing class D
415 beta-lactamases in *Acinetobacter* spp. other than *Acinetobacter baumannii*. Int J
416 Antimicrob Agents 2014; 43: 375-77.
417 [10] Koh TH, Tan TT, Khoo CT, Ng SY, Tan TY, Hsu LY, *et al. Acinetobacter*
418 *calcoaceticus-Acinetobacter baumannii* complex species in clinical specimens in
419 Singapore. Epidemiol Infect 2012; 140: 535-38.
420 [11] Wisplinghoff H, Paulus T, Lugenheim M, Stefanik D, Higgins PG, Edmond MB,
421 *et al.* Nosocomial bloodstream infections due to *Acinetobacter baumannii*,
422 *Acinetobacter pittii* and *Acinetobacter nosocomialis* in the United States. J Infect 2012;
423 64: 282-90.
424 [12] Nemec A, Krizova L, Maixnerova M, Sedo O, Brisse S, Higgins PG.
425 *Acinetobacter seifertii* sp. nov., a member of the *Acinetobacter calcoaceticus-*
426 *Acinetobacter baumannii* complex isolated from human clinical specimens. Int J Syst
427 Evol Microbiol 2015; 65: 934-42.
428 [13] Alvarez-Buylla A, Culebras E, Picazo JJ. Identification of *Acinetobacter*
429 species: is Bruker biotyper MALDI-TOF mass spectrometry a good alternative to
430 molecular techniques? Infect Genet Evol 2012; 12: 345-49.
431 [14] Hsueh PR, Kuo LC, Chang TC, Lee TF, Teng SH, Chuang YC, *et al.* Evaluation
432 of the Bruker Biotyper matrix-assisted laser desorption ionization-time of flight mass
433 spectrometry system for identification of blood isolates of *Acinetobacter* species. J Clin
434 Microbiol 2014; 52: 3095-100.

435 [15] Camoez M, Sierra JM, Domínguez MA, Ferrer-Navarro M, Vila J, Roca I.
436 Automated categorization of methicillin-resistant *Staphylococcus aureus* clinical
437 isolates into different clonal complexes by MALDI-TOF mass spectrometry. Clin
438 Microbiol Infect 2016; 22: 161 e1-7.
439 [16] Sedo O, Nemec A, Krizova L, Kacalova M, Zdrahal Z. Improvement of
440 MALDI-TOF MS profiling for the differentiation of species within the *Acinetobacter*
441 *calcoaceticus-Acinetobacter baumannii* complex. Syst Appl Microbiol 2013; 36: 572-
442 78.
443 [17] Sousa C, Botelho J, Silva L, Grosso F, Nemec A, Lopes J, *et al.* MALDI-TOF
444 MS and chemometric based identification of the *Acinetobacter calcoaceticus-*
445 *Acinetobacter baumannii* complex species. Int J Med Microbiol 2014; 304: 669-77.
446 [18] Toh BE, Paterson DL, Kamolvit W, Zowawi H, Kvaskoff D, Sidjabat H, *et al.*
447 Species identification within *Acinetobacter calcoaceticus-baumannii* complex using
448 MALDI-TOF MS. J Microbiol Methods 2015; 118: 128-32.
449 [19] Bohme K, Fernandez-No IC, Barros-Velazquez J, Gallardo JM, Calo-Mata P,
450 Canas B. Species differentiation of seafood spoilage and pathogenic gram-negative
451 bacteria by MALDI-TOF mass fingerprinting. J Proteome Res 2010; 9: 3169-83.
452 [20] Chuang YC, Sheng WH, Li SY, Lin YC, Wang JT, Chen YC, *et al.* Influence of
453 genospecies of *Acinetobacter baumannii* complex on clinical outcomes of patients with
454 acinetobacter bacteremia. Clin Infect Dis 2011; 52: 352-60.

TABLES

TABLE 1. ClinProTools peak statistics for all the species-specific peaks

Peak number	Mass	DAve	PTTA	PWKW	PAD	*Abau*	*Anos*	*Apit*	*Adij*	*Asei*
21	3948.53	13.65	0.00344	0.0000533	< 0.000001					■
23	4069.73	4.81	0.000539	< 0.000001	< 0.000001		■			
26	4180.56	9.36	< 0.000001	< 0.000001	< 0.000001		■			
27	4194.14	13.42	< 0.000001	< 0.000001	< 0.000001					■
34	4411.95	11.99	< 0.000001	< 0.000001	< 0.000001			■		
35	4430.15	37.02	< 0.000001	< 0.000001	< 0.000001				■	
58	5747.48	137.56	< 0.000001	< 0.000001	< 0.000001	■				
59	5777.3	157.06	< 0.000001	< 0.000001	< 0.000001			■		
60	5788.92	101.76	< 0.000001	< 0.000001	< 0.000001				■	
78	6692.55	5.97	0.00000267	< 0.000001	< 0.000001			■		
80	6729.52	3.94	0.0000755	< 0.000001	< 0.000001				■	
88	7446.28	8.73	0.000984	0.0000173	< 0.000001					■
89	7893.14	61.67	0.000187	< 0.000001	< 0.000001					■
92	8135.43	33.8	< 0.000001	< 0.000001	< 0.000001		■			
96	8358.1	50.16	< 0.000001	< 0.000001	< 0.000001		■			
97	8385.13	53.54	< 0.000001	< 0.000001	< 0.000001					■
101	8821.13	43.15	< 0.000001	< 0.000001	< 0.000001			■		
102	8857.63	103.26	< 0.000001	< 0.000001	< 0.000001				■	

Peak number: correlative numbering of the peak in the average spectra; Mass: *m/z* value; DAve: difference between the maximal and the minimal average peak area/intensity of all the species; PTTA: p value of t-/analysis of variance test; PWKW: p value of Wilcoxon/Kruskal–Wallis test (preferable for non-normally distributed data); PAD: p value of Anderson–Darling test, which gives information about normal distribution (p-

value AD <0.05, non-normally distributed; p-value AD >0.05, normally distributed); *Abau*: *A. baumannii*; *Anos*: *A. nosocomialis*; *Apit*: *A. pittii*; *Adij*: *A. dijkshoorniae*; *Asei*: *A. seifertii*. Shaded boxes indicate species specificity.

TABLE 2. External validation of the supervised neural network (SNN) model and the novel mean spectra (MSPs) using the ClinProTools and the MALDI BioTyper software, respectively.

Method	*Acinetobacter* species	Nº of Isolates	Nº of spectra	Spectra classification					Sen (%)	Spe (%)	PPV (%)
				A. baumannii	*A. nosocomialis*	*A. pittii*	*A. dijkshoorniae*	*A. seifertii*			
ClinProTools	*A. baumannii*	9	90	89	1	0	0	0	98.9	100	
	A. nosocomialis	11	110	0	110	0	0	0	100	96.7	
	A. pittii	7	70	0	7	61	0	2	87.1	99.7	96.8
	A. dijkshoorniae	6	60	0	0	1	59	0	98.3	100	
	A. seifertii	5	50	0	1	0	0	49	98.0	99.4	
BioTyper (Bacterial extracts)	*A. baumannii*	16	160	158	2	0	0	0	98.8	99.8	99.6
	A. nosocomialis	24	240	1	239	0	0	0	99.6	99.6	
	A. pittii	15	150	0	0	150	0	0	100	100	
	A. dijkshoorniae	12	120	0	0	0	120	0	100	100	
	A. seifertii	11	110	0	0	0	0	110	100	100	
BioTyper (Direct colonies)	*A. baumannii*	16	64	63	1	0	0	0	98.4	98.0	96.8
	A. nosocomialis	24	96	5	91	0	0	0	94.8	99.5	
	A. pittii	15	60	0	0	60	0	0	100	98.4	
	A. dijkshoorniae	12	48	0	0	4	44	0	91.7	100	
	A. seifertii	11	44	0	0	0	0	44	100	100	

Sen (%), sensitivity (%); Spe (%), specificity (%); PPV, positive predictive value

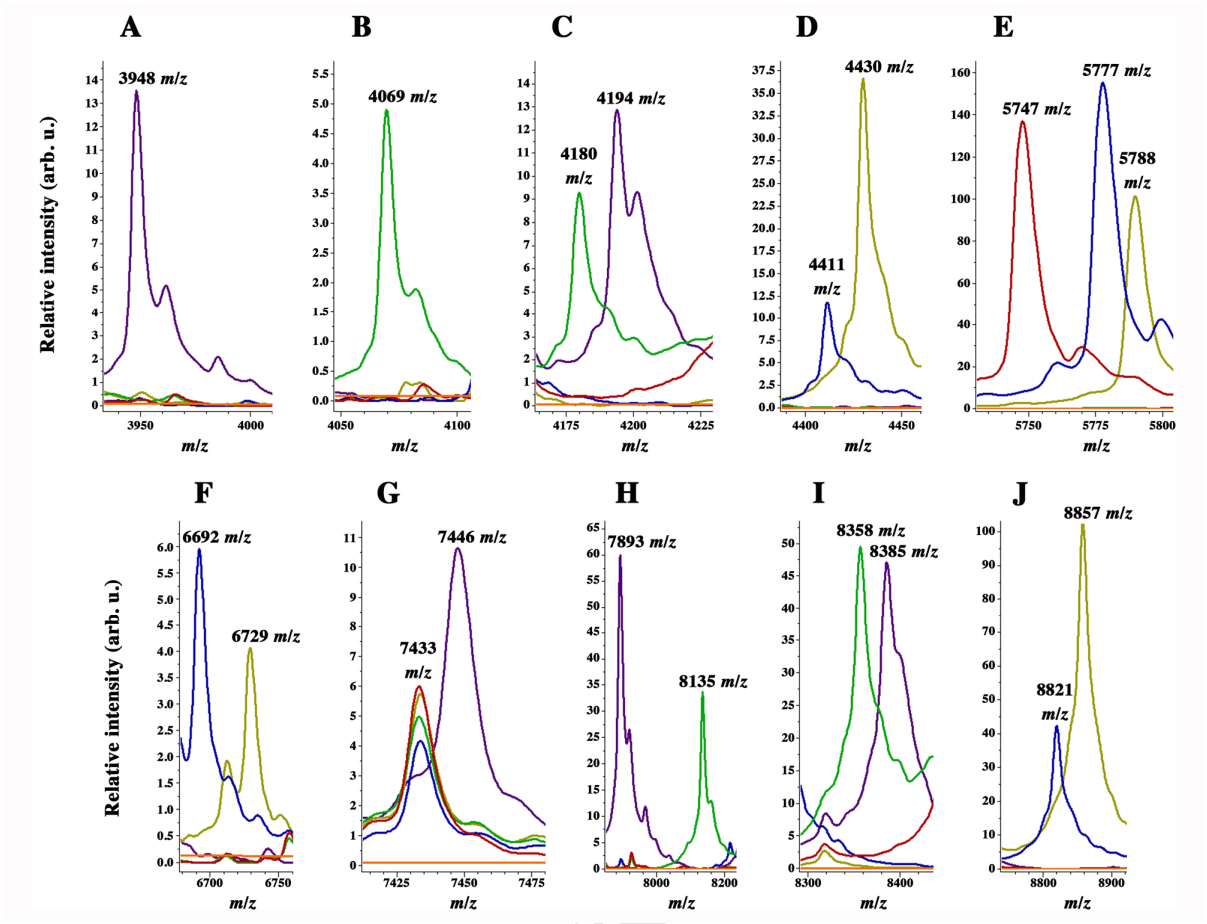
FIGURE CAPTIONS

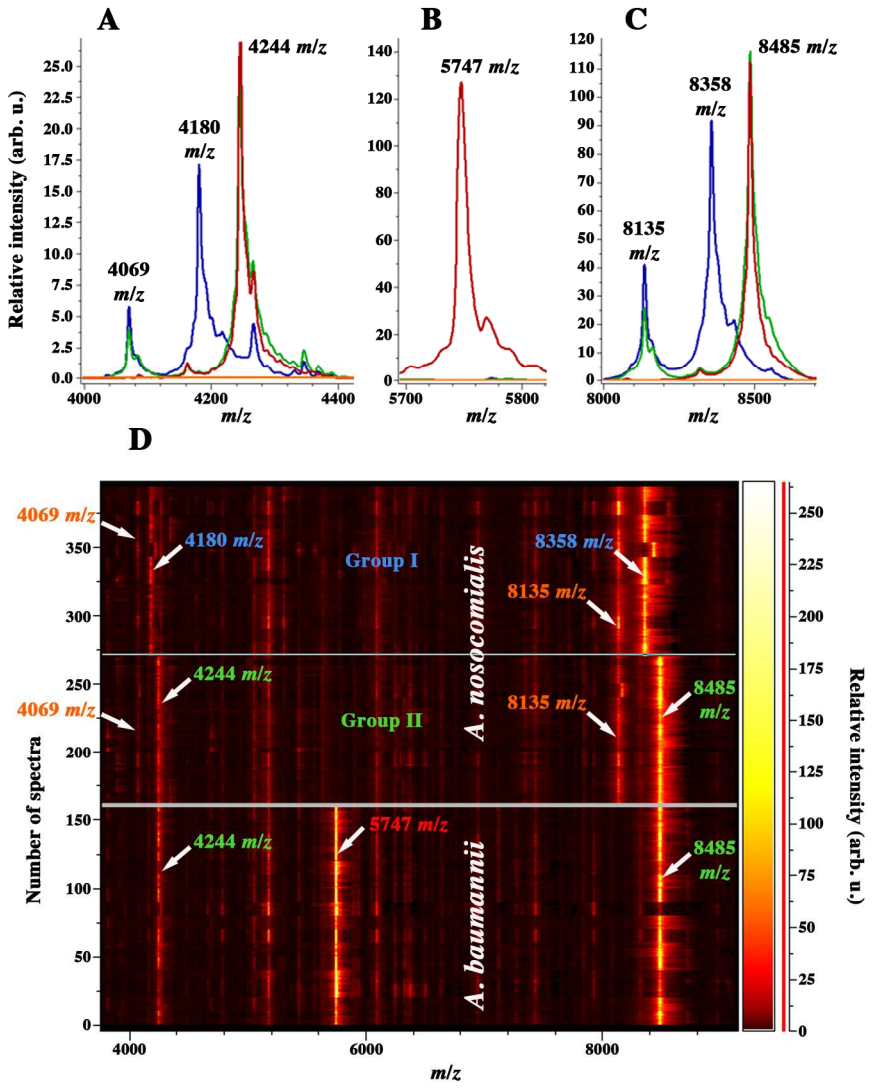
FIGURE 1. MALDI-TOF/MS averaged spectra plots from isolates of the Ab group showing specific peaks for: *A. baumannii* (red), *A. nosocomialis* (green), *A. pittii* (blue), A. *dijkshoorniae* (yellow) and A. *seifertii* (purple). The background noise signal is shown in orange. The x-axis shows the *m/z* values and the y-axis indicates the intensities of the peaks expressed in arbitrary intensity units. Peaks are ordered from left to right as A-J according to their ascending *m/z* values.

FIGURE 2. Spectral analysis of *A. nosocomialis* and *A. baumannii* isolates. *A. nosocomialis* isolates are clustered into groups I and II according to BioTyper results. **(A, B and C)** Averaged spectra plots for all the spectra included within each group. *A. baumannii* (red), *A. nosocomialis* group I (blue), *A. nosocomialis* group II (green), background noise signal (orange). The x-axis shows the *m/z* values and the y-axis indicates the intensities of the peaks expressed in arbitrary intensity units. **(D)** Gel view representation in quadratic mode and chromatic scale of all independent spectra within the 4,000-9,000 *m/z* mass range. Each isolate is represented by 10 independent spectra. The x-axis shows the *m/z* values and the y-axis indicates the number of spectra (left) as well as intensities of the peaks expressed in arbitrary intensity units (right). Grey lines are used to separate spectra from different groups. Arrows indicate *m/z* values that are present in: *A. nosocomialis* group I and II (orange labels); only in *A. nosocomialis* group I (blue labels); in both *A. nosocomialis* group II and in *A. baumannii* (green labels); and only in *A. baumannii* (red label).

FIGURE 3. Cluster analysis of all the 78 Ab group isolates included in the study based on the concatenated partial sequences of the *cpn60*, *fusA*, *gltA*, *pyrG*, *recA*, *rplB* and

494 *rpoB* genes used for MLST under the Pasteur scheme. The partial sequences of the
495 individual genes used for MLSA can be retrieved from the PubMLST website
496 (http://pubmlst.org/abaumannii/) under the sequence type codes listed in
497 **Supplementary Table S1**. Phylogenetic trees were constructed using the neighbour-
498 joining method with genetic distances computed by Kimura's two-parameter model
499 (Kimura, 1980) with a bootstrap value of 1000 replicates. Bootstrap values (%) are
500 indicated above the branches. The scale bar indicates sequence divergence. **(A)**
501 Collapsed phylogenetic tree showing the monophyletic clustering of isolates from each
502 *Acinetobacter* species within the Ab group. **(B)** Expanded phylogenetic tree showing
503 the clustering of all the *A. nosocomialis* isolates. Circles (in blue) and squares (in green)
504 indicate *A. nosocomialis* isolates classified as belonging to group I (correct BioTyper
505 identification) or group II (BioTyper misidentification), respectively, using the default
506 Bruker taxonomy database. The MSP label indicates isolates that originated the 5
507 reference spectra (MSPs) for *A. nosocomialis* currently included in the default Bruker
508 taxonomy database.
509
510 **FIGURE 4.** MSP dendrogram containing all the MALDI-TOF/MS specific signatures
511 of isolates from the *Acinetobacter calcoaceticus-Acinetobacter baumannii* complex
512 included within the default Bruker taxonomy database as well as specific signatures for
513 *A. dijkshoorniae, A. seifertii* and *A. nosocomialis* created in this study. Distance values
514 are relative and normalised to a maximal value of 1,000. The novel MSPs created in this
515 study are labelled with a single asterisk (*****). MSPs from isolates that failed to cluster
516 with their corresponding *Acinetobacter* species are labelled with a double asterisk (******).
517 T: Type strain.
518




A
B
C
D
Relative intensity (arb. u.)
4069 m/z
4180 m/z
4244 m/z
5747 m/z
8135 m/z
8358 m/z
8485 m/z
m/z
Group I
Group II
A. nosocomialis
A. baumannii
Number of spectra
Relative intensity (arb. u.)