Original article

Quality of MALDI-TOF mass spectra in routine diagnostics: results from an international external quality assessment including 36 laboratories from 12 countries using 47 challenging bacterial strains

Aline Cuénod [1, 2, *, †], Martina Aerni [3, †], Claudia Bagutti [4, †], Banu Bayraktar [5, †], Efe Serkan Boz [6, †], Cynthia Beisert Carneiro [7, †], Carlo Casanova [8, †], Alix T. Coste [9, †], Peter Damborg [10, †], Dirk W. van Dam [11, †], Mehmet Demirci [12, †, §], Pavel Drevinek [13, †, §], Olivier Dubuis [14, †], José Fernandez [15, †], Gilbert Greub [9, †, §], Jaroslav Hrabak [16, †, §], Gülen Hürkal Yiğitler [5, †], Jakub Hurych [13, †], Thøger Gorm Jensen [17, †], Géraldine Jost [18, †], Greetje A. Kampinga [19, †], Sonja Kittl [20, †], Christine Lammens [21, †], Claudia Lang [14, †], Reto Lienhard [22, †, §], Julie Logan [23, †], Carola Maffioli [24, †], Ivana Mareković [25, †, §], Matthias Marschal [26, †], Jacob Moran-Gilad [27, †, §], Oliver Nolte [28, †], Michael Oberle [29, †], Michael Pedersen [30, †], Valentin Pflüger [31, †], Sigrid Pranghofer [32, †], Julia Reichl [33, †], Rob J. Rentenaar [34, †], Arnaud Riat [15, †], Belén Rodríguez-Sánchez [35, †, §], Camille Schilt [22, †], Ann-Kathrin Schlotterbeck [1, †], Jacques Schrenzel [15, †, §], Shani Troib [27, †], Elise Willems [36, †], Mandy Wootton [37, †], Dominik Ziegler [38, †], Adrian Egli [1, 2, †, §], for the ESGMD study group

[1] Applied Microbiology Research, Department of Biomedicine, University of Basel, Basel, Switzerland
[2] Division of Clinical Bacteriology and Mycology, University Hospital Basel, Basel, Switzerland
[3] Labor Team W, Goldach, Switzerland
[4] State Laboratory Basel-Stadt, Basel, Switzerland
[5] University of Health Sciences, Sisli Hamidiye Etfal Teaching and Research Hospital, Istanbul, Turkey
[6] Department of Medical Microbiology, University of Health Sciences, Haydarpasa Numune Teaching and Research Hospital, Istanbul, Turkey
[7] University Hospital Freiburg, Freiburg im Breisgau, Germany
[8] Institute for Infectious Diseases, University of Bern, Bern, Switzerland
[9] Institute of Microbiology, University Hospital Lausanne, Lausanne, Switzerland
[10] Department of Veterinary and Animal Sciences, University of Copenhagen, Frederiksberg, Denmark
[11] Zuyderland MC, Sittard, the Netherlands
[12] Department of Medical Microbiology, Kirklareli University, Kirklareli, Turkey
[13] Department of Medical Microbiology, 2nd Faculty of Medicine, Charles University and Motol University Hospital, Prague, Czech Republic
[14] Viollier AG, Allschwil, Switzerland
[15] Division of Laboratory Medicine, Laboratory of Bacteriology, University Hospital of Geneva, Geneva, Switzerland
[16] Biomedical Center, Faculty of Medicine in Pilsen, Charles University, Plzen, Czech Republic
[17] Department of Clinical Microbiology, Odense University Hospital, Odense, Denmark
[18] Dianalabs, Geneva, Switzerland
[19] Department of Medical Microbiology and Infection Prevention, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands
[20] Institute of Veterinary Bacteriology, University of Bern, Bern, Switzerland
[21] Department of Medical Microbiology, University of Antwerp, Belgium
[22] ADMED Microbiologie, La Chaux de Fonds, Switzerland
[23] Reference Services Division, UK Health Security Agency, London, United Kingdom
[24] MCL Laboratories, Niederwangen, Switzerland
[25] Department of Clinical and Molecular Microbiology, University Hospital Centre Zagreb, Zagreb, Croatia
[26] Institute of Medical Microbiology and Hygiene, University of Tübingen, Tübingen, Germany
[27] School of Public Health, Ben Gurion University of the Negev and Soroka University Medical Center, Beer Sheva, Israel
[28] Center for Laboratory Medicine, St. Gallen, Switzerland
[29] Cantonal Hospital Aarau, Aarau, Switzerland

* Corresponding author: A. Cuénod, Applied Microbiology Research, Department of Biomedicine, University of Basel, Basel, Switzerland.
  E-mail address: aline.cuenod@stud.unibas.ch (A. Cuénod).
† Authors in alphabetical order.
§ European Study group on Genomics and Molecular Diagnosis (ESGMD) of the European Society of Clinical Microbiology and Infectious Disease (ESCMID), Basel, Switzerland

[30] *Department of Clinical Microbiology, Hvidovre Hospital, Hvidovre, Denmark*

[31] *Mabritec AG, Riehen, Switzerland*

[32] *Bioanalytica AG, Lucerne, Switzerland*

[33] *Austrian Agency for Health and Food Safety, Vienna, Austria*

[34] *UMC Utrecht, Utrecht, the Netherlands*

[35] *Hospital General Universitario Gregorio Maranon, Madrid, Spain*

[36] *Clinical Laboratory AZNikolaas, Sint-Niklaas, Belgium*

[37] *University Hospital of Wales, Cardiff, United Kingdom*

[38] *Eurofins Scientific AG, Schönenwerd, Switzerland*

ARTICLE INFO

ABSTRACT

Objectives: Matrix assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry (MS) is a widely used method for bacterial species identification. Incomplete databases and mass spectral quality (MSQ) still represent major challenges. Important proxies for MSQ are the number of detected marker masses, reproducibility, and measurement precision. We aimed to assess MSQs across diagnostic laboratories and the potential of simple workflow adaptations to improve it.

Methods: For baseline MSQ assessment, 47 diverse bacterial strains, which are challenging to identify by MALDI-TOF MS, were routinely measured in 36 laboratories from 12 countries, and well-defined MSQ features were used. After an intervention consisting of detailed reported feedback and instructions on how to acquire MALDI-TOF mass spectra, measurements were repeated and MSQs were compared.

Results: At baseline, we observed heterogeneous MSQ between the devices, considering the median number of marker masses detected (range = [2−25]), reproducibility between technical replicates (range = [55%−86%]), and measurement error (range = [147 parts per million (ppm)−588 ppm]). As a general trend, the spectral quality was improved after the intervention for devices, which yielded low MSQs in the baseline assessment as follows: for four out of five devices with a high measurement error, the measurement precision was improved (p-values <0.001, paired Wilcoxon test); for six out of ten devices, which detected a low number of marker masses, the number of detected marker masses increased (p-values <0.001, paired Wilcoxon test).

Discussion: We have identified simple workflow adaptations, which, to some extent, improve MSQ of poorly performing devices and should be considered by laboratories yielding a low MSQ. Improving MALDI-TOF MSQ in routine diagnostics is essential for increasing the resolution of bacterial identification by MALDI-TOF MS, which is dependent on the reproducible detection of marker masses. The heterogeneity identified in this external quality assessment (EQA) requires further study. **Aline Cuénod, Clin Microbiol Infect 2023;29:190**

Introduction

Matrix assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry (MS) is a commonly used method for microbial species identification in modern diagnostic laboratories [1−3] due to its minimal hands-on and short turnaround time, cost-efficiency, and high accuracy [4,5].

Multiple studies have shown the improved resolution gained by using marker-based analytical approaches [6−9] compared to pattern matching approaches. This insight has led to the development of marker-based databases for bacterial identification [8,10], such as the PAPMID database (Mabritec AG, Riehen, Switzerland) [11]. In such approaches, specific peaks of interest, whose presence is associated with a species [12], lineage [13], or even mobile genetic elements [14,15], are queried in the acquired mass spectrum in order to increase specificity and resolution. Many of the peaks, which can be reproducibly detected in MALDI-TOF mass spectra, correspond to protein subunits of the bacterial ribosome [16]. A high MALDI-TOF mass spectral quality (MSQ) is required in order to reproducibly detect marker peaks.

Despite the success of MALDI-TOF MS for routine microbial species identification, multiple clinically-relevant species are currently not distinguished by commonly used databases using pattern matching approaches. Possible reasons for this are that (a) the databases are incomplete, (b) the species of interest resemble closely other species in the databases, and (c) MALDI-TOF mass spectra are of low quality. We previously compiled a diverse set of 47 bacterial strains, representing 39 species and 15 genera, which are difficult to be identified at a species level for the above-mentioned reasons [17]. In this previous publication [17], we defined the following five important spectrum features as good proxies for MSQ: (a) the number of ribosomal marker peaks detected, (b) the median relative intensity of ribosomal marker peaks, (c) the sum of the intensity of all detected peaks, (d) a high measurement precision, and (e) reproducibility of peaks between technical replicates. Determining these MSQ features, we previously assessed the performance of different sample preparation protocols on different bacterial groups and consequently proposed to use the formic acid overlay protocol for unknown samples and group specific protocols for highest MSQ [17]. Whether the proposed protocols can effectively increase MSQ of these challenging strains in routine settings has yet to be evaluated.

The aim of this study was therefore to assess (a) the MSQ obtained in routine diagnostics, (b) whether there are routine practices associated with an increased MSQ, (c) whether the MSQ can be improved using the protocols proposed, and (d) compile a reference
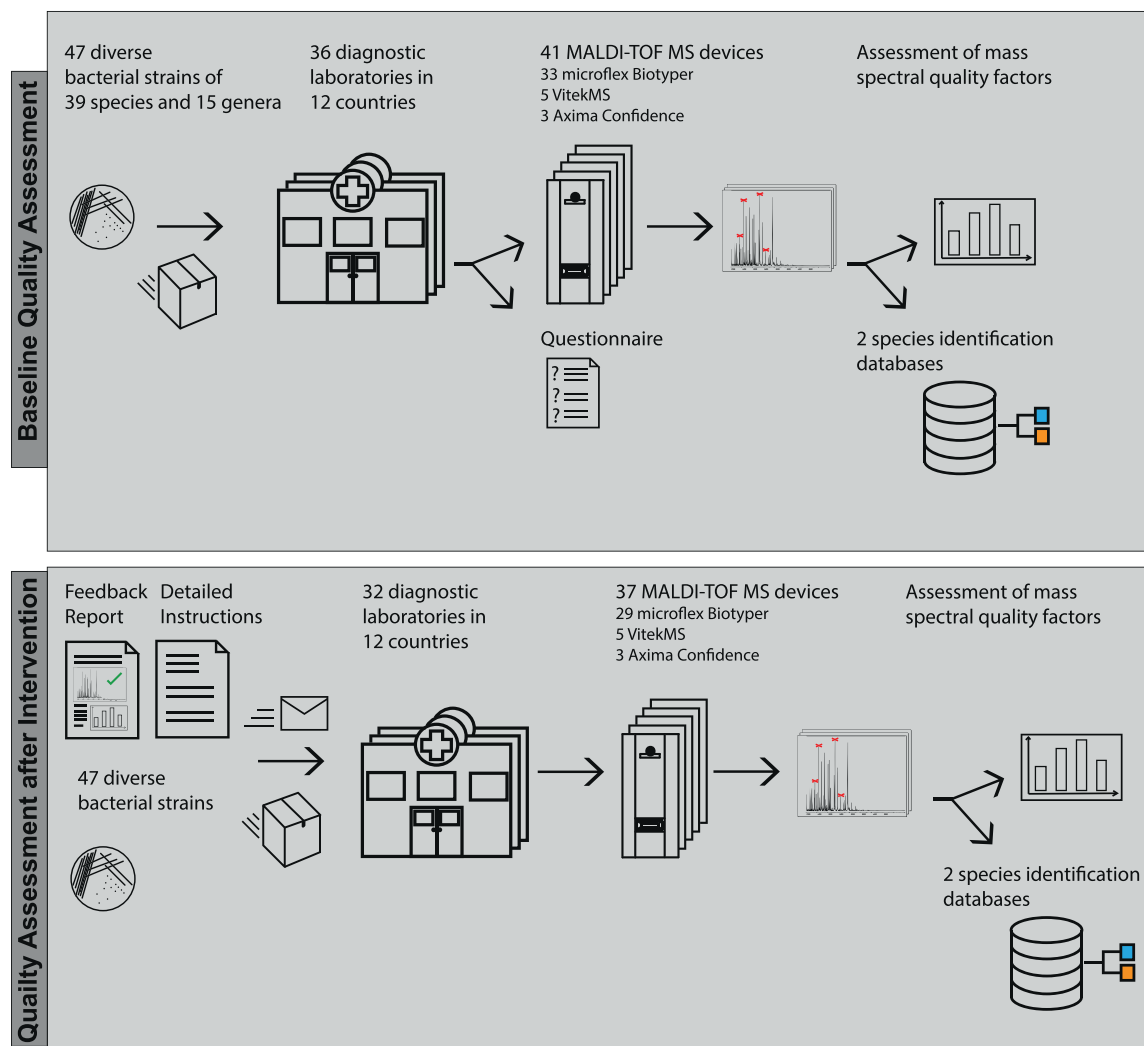
Fig. 1. Overview on the workflow of the study. The upper panel shows the baseline quality assessment including 36 participating laboratories and the lower panel shows the post-interventional quality assessment including 32 laboratories using the same bacterial strains. MALDI-TOF, matrix assisted laser desorption ionization–time of flight; MS, mass spectrometry.

dataset of MALDI-TOF mass spectra including technical replicates, matching genomic sequences, and extensive metadata.

Methods

Design of the external quality assessment (EQA)

Fig. 1 provides an overview over the workflow of this study.

Bacterial strains

The bacterial strains used in this study have previously been described [17] and their whole genome sequences (see Supplementary material, Table S1), as well as the previously predicted masses of the ribosomal subunits, are publicly available (https://osf.io/ksz7r/). See Supplementary Methods for more detail on these strains and the participating laboratories.

Baseline MALDI-TOF MSQ assessment

The participating laboratories were asked to culture the bacterial isolates and acquire MALDI-TOF mass spectra according to their

routine diagnostic procedures, which may vary between the laboratories. Each laboratory was asked to fill out a questionnaire on routine laboratory practice.

Intervention

Each participating laboratory received a feedback report on the MALDI-TOF mass spectra acquired for the baseline quality assessment (example in Supplementary File 1) and instructions on how to acquire MALDI-TOF mass spectra in subsequent measurements of the same strains, aiming to improve the MSQ using a standardised approach (see Supplementary File 2).

We provided the following two different sets of protocols: (a) a simple 'generic protocol' (i.e., 'formic acid overlay') for all samples and (b) group-specific sample preparation protocols, aiming at highest MSQ [17] (see Supplementary File 2 for more detail).

MALDI-TOF mass spectra processing

Peaks were picked from raw spectra using default settings by the software included in the microflex Biotyper or the VitekMS/Axima Confidence system (see Supplementary Methods for more detail).

The raw data acquired on each device, the processed peak list and the species identification results of all databases used can be accessed via the Open Science Foundation (https://osf.io/ae2nk/).

We queried each spectrum for the following features to assess the MSQ: (a) the number of ribosomal marker peaks detected, (b) the median relative intensity of ribosomal marker peaks, (c) the sum of the intensity of all detected peaks, (d) the measurement precision, and (e) reproducibility of peaks between technical replicates. As factors (a) through (c) often correlate [17], we have focused on factors (a), (d), and (e) in the main text and figures of this study (see Supplementary Methods).

Scripts used for spectra evaluation and data visualization can be accessed via GitHub (https://github.com/appliedmicrobiology research/MALDI-TOF-MS-EQA).

Databases used for species identification

Each spectrum acquired on a Bruker device was compared to the MALDI Biotyper database (MALDI Biotyper Compass Library, Revision E (v8.0, 8468 MSP, RUO, Bruker Daltonics, Bremen, Germany). Spectra acquired on a Axima Confidence or VitekMS device were analysed with the VitekMS database (v3.2, bioMérieux, Marcy-l'Étoile, France).

Furthermore, we compared each spectrum to a ribosomal marker-based database, either PAPMID^{TM} or PAPMID^{TM} subtyping modules (both Mabritec AG, Riehen, Switzerland, henceforward be referred to as PAPMID).

In the main text of this manuscript, we report the species identification by the PAPDIM^{TM} database, as this database (a) allows species identification from spectra acquired on devices of different manufacturers and (b) includes all species represented by our strain set.

To evaluate species identification, we classify the results of the PAPMID^{TM} database into the following accuracy categories: (a) the correct species unambiguously receives the highest score ('correct identification'), (b) the correct species and other species receive the highest score i.e., the identification is correct but ambiguous ('correct multi-species identification'), (c) the score is below the identification threshold and no species identification is possible ('no identification possible'), and (d) the identified species is unambiguously wrong ('wrong identification').

More details about database scores and their interpretations can be found in the Supplementary Methods.

Statistical analysis

We used paired Wilcoxon rank tests when comparing spectra acquired from the same strains and excluded spectra of strains which were missing in one of the sets of interest. We used unpaired Wilcoxon rank tests (Mann Whitney U tests) when comparing spectra acquired from different strains. The nomenclature 'median (lower bound of the interquartile range (IQR), upper bound of the IQR)' was used when referring to data in the running text throughout the study.

All analyses were performed in R (v4.0.3) using the ggpubr (v4.0), the rstatix (v0.7) package, and visualised using ggplot2 (v3.3.5) [20].

Results

Access to data

All MALDI-TOF mass spectra acquired for this study, including peaklists and species identification from three different databases can be downloaded from (https://osf.io/ae2nk/). The filled-out questionnaire (see Supplementary material, Table S3) provides valuable metadata. The bacterial strains included in this study have previously been whole genome sequenced and the raw reads are publicly available (see Supplementary material, Table S1).

Heterogeneity in MSQ across diagnostic laboratories

For the baseline quality assessment, we received 5035 spectra measured on 41 devices from the 36 participating laboratories. We observed differences between the devices in MSQ considering the number of marker masses detected (e.g., device 7: median = 25; IQR = [20, 28]; device 32: median = 5; IQR = [3, 14]) (Fig. 2A).

The heterogeneity of MSQ was reflected in varying accuracy in species identification (see Supplementary material, Table S4). Over all bacterial strains and using a marker-based species identification, the fraction of spectra, which were correctly and uniquely identified to the species level, ranged from 22.5% (18/80 spectra, device 9) to 78.2% (147/188 spectra, device 35). We observed no difference in MSQ between the different MALDI-TOF MS manufacturers and an increasing accuracy of species identification with increasing MSQ (Fig. 2A).

The MSQ differed between bacterial groups with a lower MSQ observed for spectra of gram-positive isolates compared to spectra of gram-negative isolates (marker masses detected: median = 16; IQR = [12, 20] vs. median = 18; IQR = [15, 22]; p <0.0001) (Fig. 2B).

We observed a four-fold difference in measurement error when comparing the most and least precise measurements (device 11: 147 parts per million (ppm), IQR, 109 ppm−192 ppm vs. device 41: 588 ppm, IQR, 533 ppm−631 ppm) (Fig. 2C).

These differences in MSQ are mainly represented by a few participating laboratories, while the data from most laboratories cluster around the overall median = 16; IQR = [13, 19] marker masses detected and 280 ppm, IQR, 177ppm−426 ppm in measurement error.

Routine laboratory practices are associated with MSQ

Every participating laboratory filled out a questionnaire on laboratory practices (see Supplementary material, Table S3). As in each laboratory different combinations of practices apply, some of which are not reflected in this questionnaire, we cannot identify causative factors of laboratory practices on spectral quality. However, we observed that certain practices correlated with improved MSQ represented by an increased number of ribosomal subunits detected detailed as follows: (a) acquisition of spectra on steel targets compared to disposable targets (median = 17; IQR = [13, 20] vs. median = 13; IQR = [10, 16]; p <0.0001), (b) cleaning steel target plates with 'methanol-acetone' protocol compared to other cleaning protocols (median = 23; IQR= [18, 26] vs. median = 16; IQR = [12, 18]; p <0.0001), (c) regular hardware services by the MALDI-TOF MS provider (median = 17; IQR = [13, 20] vs. median =15; IQR = [11, 18]; p <0.0001), (d) working with a MALDI-TOF MS workstation (i.e., for a certain period, one or more member of staff are responsible for all MALDI-TOF MS measurements) (median = 17; IQR = [14, 21] vs. median = 16; IQR = [12, 18]; p <0.0001), (e) replacing the matrix solution after 7 or less days (median = 17; IQR = [13, 20] vs. median = 15; IQR = [11, 17]; p <0.0001), and (f) subculturing isolates on agar plates after defreezing, or culturing the isolates on agar plates from the ESwab transport medium, compared to strains which were measured directly after culturing on agar plates from frozen stocks (median = 17; IQR = [13, 20] vs. median = 11; IQR = [6, 15.8]; p <0.0001) (see Supplementary material, Fig. S3) (Fig. 3).
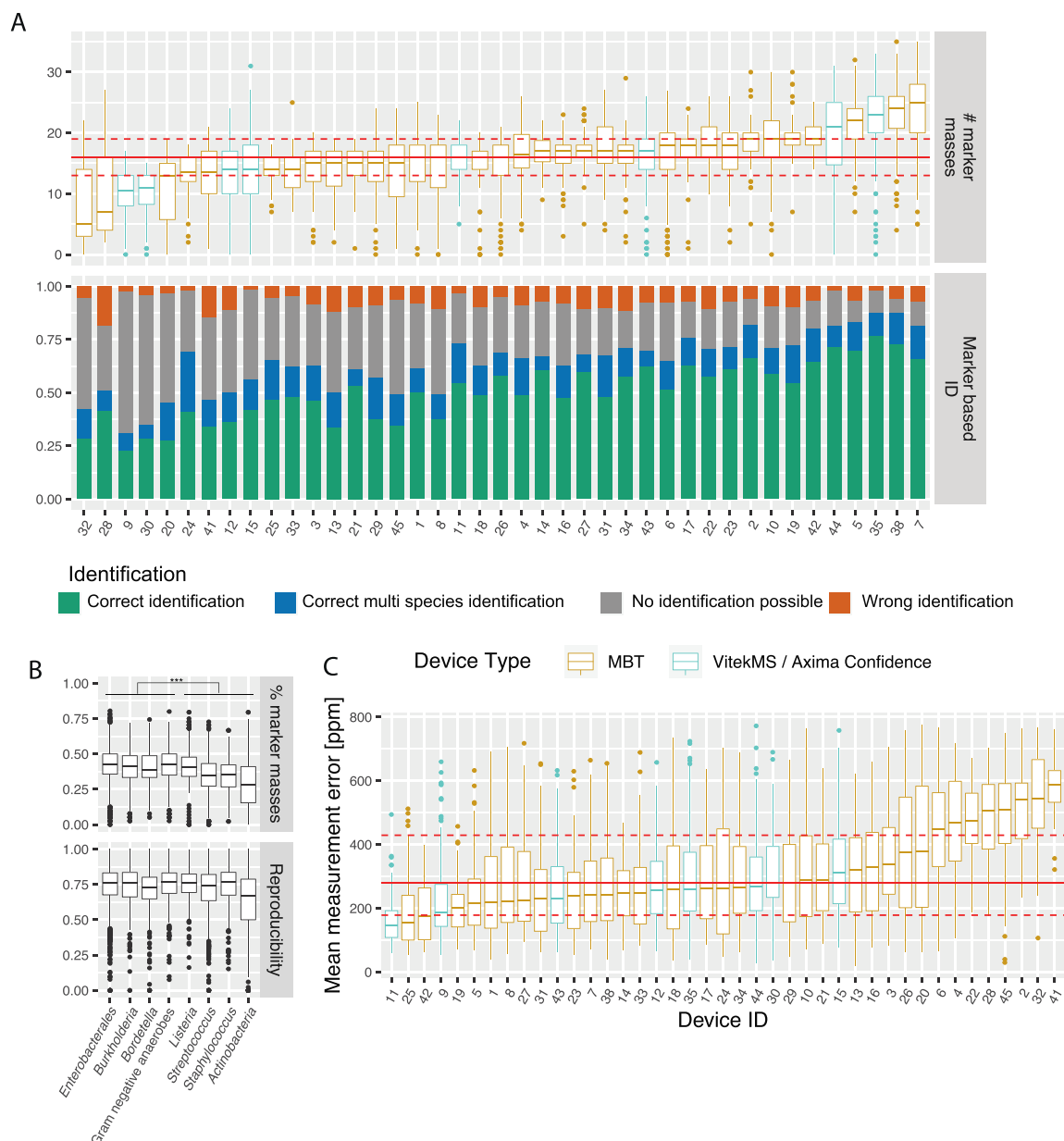
Fig. 2. (A) Number of ribosomal marker masses detected in matrix assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectra acquired on 41 devices (upper row) and the evaluation of the species identification results using a marker-based approach (lower row). (B) Relative number of marker masses detected per phylogenetic group (upper row) and reproducibility between technical replicates (lower row). (C) Measurement errors of the different devices. Boxplots: The middle line corresponds to the median, the lower and the upper hinge depict the first and the third quartile, whereas the whiskers extend from the hinge no further than 1.5 times the interquartile range. Data points beyond this range are depicted as individual points. MBT, microflex biotyper.

Effect of standardized protocols on MALDI-TOF MSQ

Calibration

We asked all participating laboratories to calibrate the devices (i.e., mass-axis calibration) before acquiring the second set of MALDI-TOF mass spectra [21]. In the calibration process, the time of flight of proteins with known mass is measured. From this, the conversion from time of flight to mass is calculated and reset for the following measurements. The measured time of flight of a protein can change with external factors such as (a) the temperature and thus the length of the flight tube, (b) the thickness of the sample, and (c) the curvature of the target. Compared to spectra acquired for the baseline quality assessment, the measurement error was significantly lower for the spectra acquired in this second round in 14 of 36 devices, no significant change was observed in 11 of 36, and a significant increase in measurement error was observed on 11 of 36 devices. When focusing on devices, at baseline yielding a measurement error above 500 ppm, we recorded a significant decrease in measurement error in 4 of 5 cases (Fig. 4).

Comparing sample preparation protocols per device

We observed an improved MSQ represented by an increased detection of marker masses using the formic acid overlay protocol compared to spectra acquired for baseline quality assessment in 10 out of 36 devices, and a decrease in 17 out of 36 devices. In 9 out of 36 devices there was no significant change. Of the devices for which the median number of marker masses was lower than 15 for
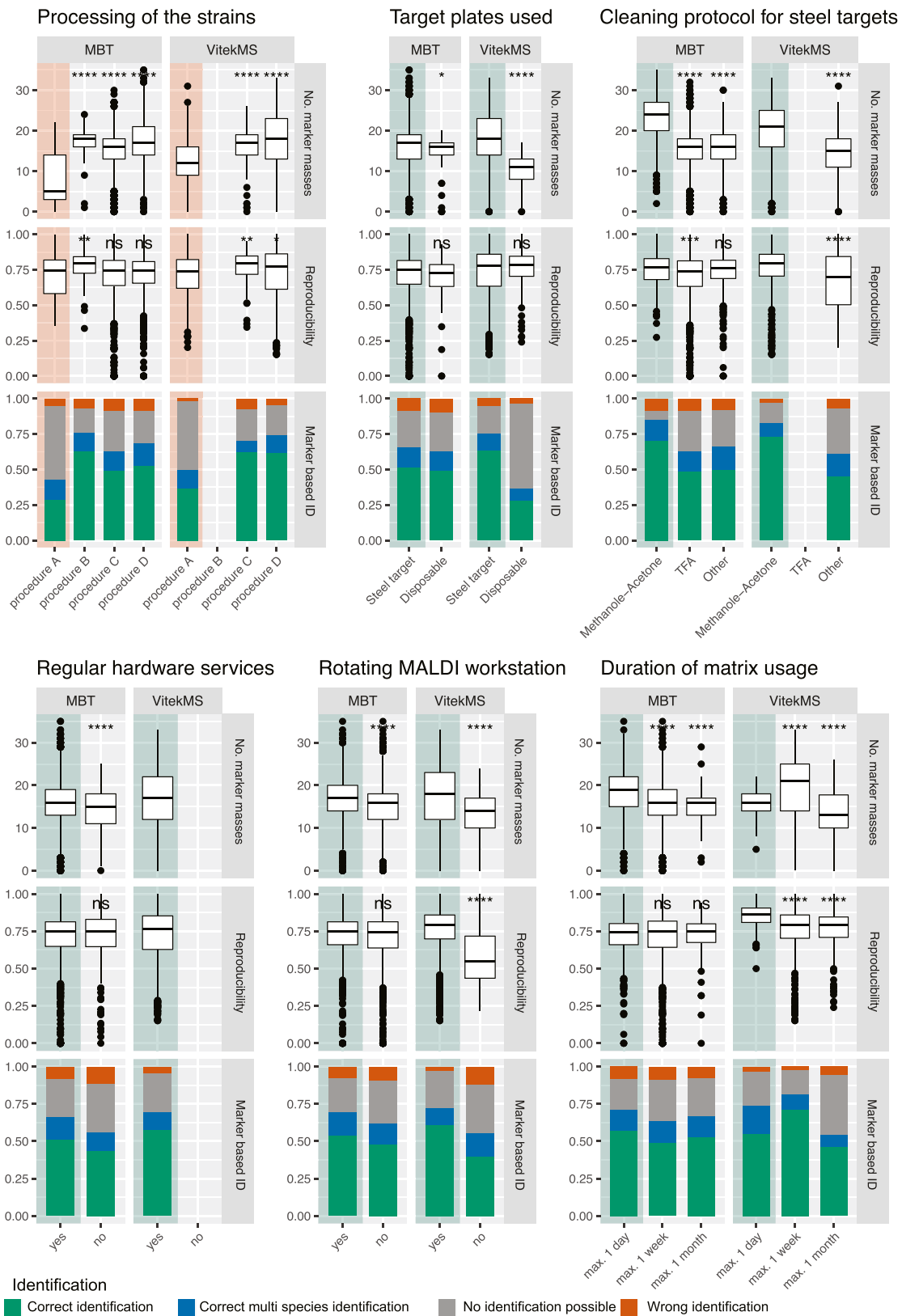
Fig. 3. Mass spectra quality features (a) number of detected marker masses and (b) technical reproducibility of A: strains processed in laboratories using different culturing procedures (procedure A: streaked out from frozen stock; procedure B: streaked out from frozen stock, subcultured once; procedure C: streaked out from ESwab; procedure D: Streaked out from ESwab, subcultured once) (left column), using different target plates (middle column) and using varying cleaning protocols (right column). B: performing hardware services or not (left column), working with a MALDI workstation or not (middle column) and keeping the matrix for varying time in the workflow (right column). ***p <0.001, unpaired Wilcoxon-rank test. Boxplots: The middle line corresponds to the median, the lower and the upper hinge depict the first and the third quartile, whereas the whiskers extend from the hinge no further than 1.5 times the interquartile range. Data points beyond this range are depicted as individual points. MALDI, matrix assisted laser desorption ionization; MBT, microflex Biotyper; VitekMS, includes VitekMS and Shimadzu devices.

Fig. 4. Measurement errors of spectra acquired in the baseline quality assessment (white) and after the intervention, which includes calibrating the device before spectra acquisition (blue). Devices are ordered, according to the median measurement errors recorded for spectra acquired for baseline quality assessment. Statistical comparisons performed using paired Wilcoxon rank tests. Boxplots: The middle line corresponds to the median, the lower and the upper hinge depict the first and the third quartile, whereas the whiskers extend from the hinge no further than 1.5 times the interquartile range. Data points beyond this range are depicted as individual points. ppm, parts per million.

baseline acquired spectra, we observed an increase of detected marker masses in 6 out of 10 devices and a decrease in 1 out of 10 devices.

Comparing the group-specific protocols to the formic acid overlay protocol, we observed an increase in the number of marker masses detected in 7 out of 34 devices, and a decrease in 18 out of 34 devices. In 9 out of 34 devices there was no significant change (Fig. 5).

Sample preparation protocols have varying effects on different bacterial groups

When comparing spectra acquired with the formic acid overlay protocol to the routinely acquired spectra per bacterial group, we surprisingly observed an increase of marker masses only in 1 out of 9 bacterial groups, namely *Staphylococcus* (Fig. 5B).

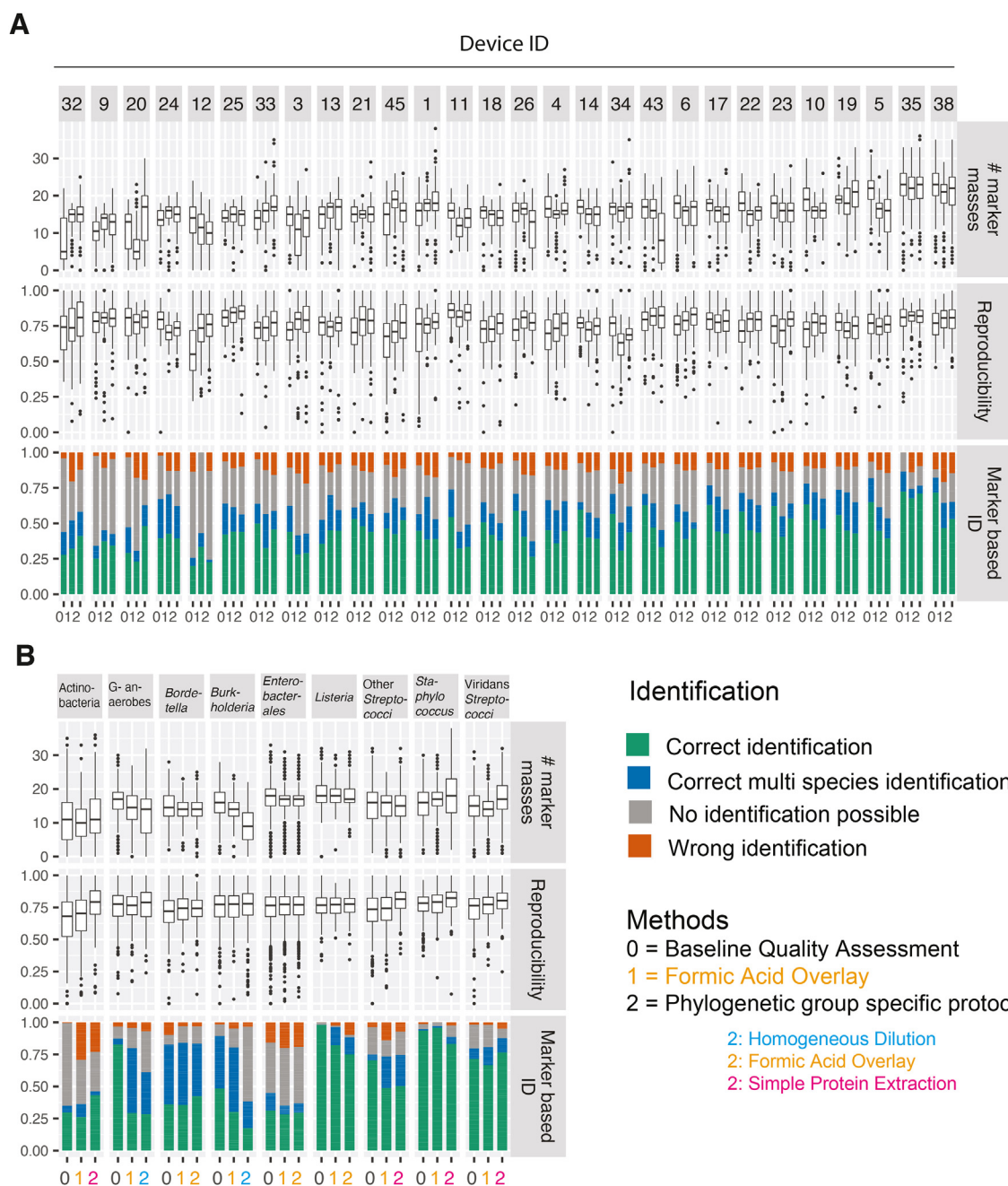Overall, we observed a negative effect from the group-specific protocols compared to the formic acid overlay protocol, which is mainly driven by two of the bacterial groups—*Burkholderia* and gram-negative anaerobes. For these two groups, the phylogenetic group-specific protocol required diluting the samples homogeneously in a buffer solution.

We observed a positive effect of using the simple protein extraction protocols compared to the formic acid overlay protocol for viridans streptococci (number of marker masses detected: median = 17; IQR = [13.75, 20] vs. median = 14; IQR = [12, 17]; p <0.0001), *Staphylococcus* (number of marker masses detected: median = 18; IQR = [13, 23] vs. median = 17; IQR = [15, 19]; p = 0.002), and *Actinobacteria* (number of phylogenetic marker masses identified: median = 12; IQR = [7, 18] vs. median = 11; IQR = [7, 14]; p = 0.03, p <0.0001).

We observed a general trend of the accuracy and resolution of a marker-based species identification following the number of ribosomal marker masses detected. However, there are exceptions to this trend such as staphylococci spectra for which we observed a higher number of non-identifiable spectra with a higher median number of ribosomal subunits detected.

Discussion

In this external quality assessment (EQA), we systematically compared the MALDI-TOF MSQ between 36 routine diagnostic laboratories, using previously defined mass spectral features. EQAs on the use of MALDI-TOF MS for microbial species identification have previously been reported [22], comparing the ability of diagnostic laboratories to identify a defined set of bacterial strains using MALDI-TOF MS. As the identification results are influenced by the reference

database and the MSQ, it is not possible to disentangle these two factors. We [17] and others [23] have previously shown how sample preparation adaptations can improve MSQ. Previous studies examining MALDI-TOF MSQ have been performed on a single device [23,24]. In this study, we assessed whether sample preparation protocols, which yielded high quality MALDI-TOF mass spectra in our hands, can increase MSQ in routine diagnostic laboratories. We thereby compiled a comprehensive dataset of MALDI-TOF mass spectra with up to 250 technical replicates per bacterial strain, with extensive metadata and matching genomic sequences being publicly available.

For the baseline quality assessment, we asked the participating laboratories to culture and measure the strains as they would do in their diagnostic workflow. This makes disentangling methodological effects difficult but reflects diagnostic reality. The spectra quality observed from these measurements might differ from the spectra quality observed in routine diagnostics, for the following reasons: (a) the participating laboratories knew beforehand that the quality would be assessed, which might have biased the participants towards putting more effort in these measurements (e.g., by repeating measurements); (b) these strains were shipped using ESwab transport media and were not cultured directly from patient material; (c) it was indicated to grow all strains on standard blood agar plate, whereas in routine diagnostics bacterial colonies might be picked from other media; and (iv) the samples were processed outside of the routine workflow and the unusual situation could have decreased MSQ.

We found a notable heterogeneity between measurements performed on different devices, which was driven by a few, poorly, and highly performing devices. The fact that the MSQ from most devices clustered around the overall median highlights the robustness of the method. When comparing spectra acquired using our suggested protocols to baseline acquired spectra, we found a positive effect for devices performing poorly at baseline, whereas the effect was often non-significant or even negative for devices performing well in the baseline quality assessment.

Hardware factors such as (a) the sort and age of the laser, (b) the cleanliness of the ion source, and (c) the tension of the detector might have an effect on MSQ and were not considered, which is a limitation of this study. As previously described [23], the simple protein extraction protocol improves MSQ for gram-positive strains, however the observed effects were modest. For *Burkholderia* and gram-negative anaerobes, we proposed to prepare homogeneous dilutions of the samples [17], which did not perform well when tested by the participating laboratories. We hypothesise that this might either be the result of (a) differing amounts of

Fig. 5. Effect of different sample preparation protocols on matrix assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectral quality (MSQ) of spectra acquired on 28 MALDI-TOF mass spectrometry (MS) devices (devices on which not all bacterial groups were measured with all three protocols were excluded from this graph). (A) Number of marker masses detected (upper row), reproducibility between technical replicates (middle row), and evaluation of a marker-based species identification (lower row) for spectra acquired with different methods and on different devices. Devices are ordered according to the number of marker masses recorded in spectra acquired for the baseline quality assessment. (B) Number of marker masses detected (upper row), reproducibility between technical replicates (middle row), and evaluation of a marker-based species identification (lower row) for spectra acquired with different methods and from various bacterial groups. Boxplots: The middle line corresponds to the median, the lower and the upper hinge depict the first and the third quartile, whereas the whiskers extend from the hinge no further than 1.5 times the interquartile range. Data points beyond this range are depicted as individual points.

bacterial inoculum used, as this was not indicated precisely enough in the instructions document or (b) differing sensitivities of the MALDI-TOF MS devices used.

We observed the accuracy and resolution of a marker-based species identification mainly following the number of detected ribosomal subunits. There are exceptions to this trend, such as *Staphylococcus aureus* complex mass spectra whose species identification did not improve with a higher median of phylogenetic

marker masses being detected. Possible explanations for this could be (a) the larger scatter of marker masses (i.e., for a larger part of the spectra the discriminatory marker masses were missing) and (b) these spectra were particularly noisy, which led to more false positive marker masses.

Based on the data analysed in this study, we suggest the following practices to be implemented in routine diagnostics: (a) regular assessment of MSQ in diagnostic laboratories, internally

(e.g., weekly), as well as externally (e.g., bi-yearly); (b) frequent calibration of the devices using well defined mass-standards; (c) usage of group specific protocols, whenever routine sample preparation does not yield satisfactory MSQ.

A frequent MSQ assessment could help to notice a drop in MSQ in a timely manner. Depending on the supposed cause for the decreased MSQ, possible responses could be to (a) adjust the devices hardware settings (e.g., the tension of the detector) or (b) refresh the personnel's skills for sample preparation.

We have identified simple workflow adaptations which improve MSQ of poorly performing devices. These implementations could increase the number of reproducibly detected marker masses in routine diagnostics. More reproducibly detected peaks increase the feasibility for MALDI-TOF MS-based typing and might improve the early recognition of spreading clones. The heterogeneity found in this EQA deserves further study in order to optimise MALDI-TOF MS-based routine identification in clinical laboratories.

Author contributions

Planning and conceptualization of the study: A.C. and A.E.; data acquisition: all authors; data analysis: A.C.; writing of the manuscript: A.C. and A.E.; giving relevant feedback throughout the project and on the manuscript: all authors.

Transparency declaration

V.P. is an employee of the company Mabritec AG, Riehen, Switzerland, which commercializes ribosomal marker-based approaches in MALDI-TOF MS data analyses for identification of microorganisms and develops the PAPMID database. The remaining authors declare that they have no competing interests.

Funding

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cmi.2022.05.017.

References

[1] Angeletti S. Matrix assisted laser desorption time of flight mass spectrometry (MALDI-TOF MS) in clinical microbiology. J Microbiol Meth 2017;138:20–9. https://doi.org/10.1016/j.mimet.2016.09.003.

[2] Kostrzewa M. Application of the MALDI Biotyper to clinical microbiology: progress and potential. Expert Rev Proteomics 2018;15:193–202. https://doi.org/10.1080/14789450.2018.1438193.

[3] Welker M, Van Belkum A, Girard V, Charrier JP, Pincus D. An update on the routine application of MALDI-TOF MS in clinical microbiology. Expert Rev Proteomics 2019;16:695–710. https://doi.org/10.1080/14789450.2019.1645603.

[4] Clark AE, Kaleta EJ, Arora A, Wolk DM. Matrix-assisted laser desorption ionization–time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. Clin Microbiol Rev 2013;26:547–603. https://doi.org/10.1128/CMR.00072-12.

[5] Feucherolles M, Poppert S, Utzinger J, Becker SL. MALDI-TOF mass spectrometry as a diagnostic tool in human and veterinary helminthology: a systematic review. Parasit Vectors 2019;12:245. https://doi.org/10.1186/s13071-019-3493-9.

[6] Rothen J, Pothier JF, Foucault F, Blom J, Nanayakkara D, Li C, et al. Subspecies typing of streptococcus agalactiae based on ribosomal subunit protein mass variation by MALDI-TOF MS. Front Microbiol 2019;10:471. https://doi.org/10.3389/fmicb.2019.00471.

[7] Ziegler D, Pothier JF, Ardley J, Fossou RK, Pflüger V, de Meyer S, et al. Ribosomal protein biomarkers provide root nodule bacterial identification by MALDI-TOF MS. Appl Microbiol Biotechnol 2015;99:5547–62. https://doi.org/10.1007/s00253-015-6515-3.

[8] Bridel S, Watts SC, Judd LM, Harshegyi T, Passet V, Rodrigues C, et al. Klebsiella MALDI TypeR: a web-based tool for Klebsiella identification based on MALDI-TOF mass spectrometry. Res Microbiol 2021;172:103835. https://doi.org/10.1016/j.resmic.2021.103835.

[9] Ojima-Kato T, Yamamoto N, Suzuki M, Fukunaga T, Tamura H. Discrimination of Escherichia coli O157, O26 and O111 from other serovars by MALDI-TOF MS based on the S10-GERMS method. PLoS One 2014;9:e113458. https://doi.org/10.1371/journal.pone.0113458.

[10] Tomachewski D, Galvão CW, de Campos Júnior A, Guimarães AM, Ferreira da Rocha JC, Etto RM. Ribopeaks: a web tool for bacterial classification through m/z data from ribosomal proteins. Bioinformatics 2018;34:3058–60. https://doi.org/10.1093/bioinformatics/bty215.

[11] Kassim A, Pflüger V, Premji Z, Daubenberger C, Revathi G. Comparison of biomarker based matrix assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) and conventional methods in the identification of clinically relevant bacteria and yeast. BMC Microbiol 2017;17:128. https://doi.org/10.1186/s12866-017-1037-z.

[12] Cuénod A, Wüthrich D, Seth-Smith HMB, Ott C, Gehringer C, Foucault F, et al. Whole-genome sequence-informed MALDI-TOF MS diagnostics reveal importance of Klebsiella oxytoca group in invasive infections: a retrospective clinical study. Genome Med 2021;13:150. https://doi.org/10.1186/s13073-021-00960-5.

[13] Lafolie J, Sauget M, Cabrolier N, Hocquet D, Bertrand X. Detection of Escherichia coli sequence type 131 by matrix-assisted laser desorption ionization time-of-flight mass spectrometry: implications for infection control policies? J Hosp Infect 2015;90:208–12. https://doi.org/10.1016/j.jhin.2014.12.022.

[14] Josten M, Dischinger J, Szekat C, Reif M, Al-Sabti N, Sahl HG, et al. Identification of agr-positive methicillin-resistant Staphylococcus aureus harbouring the class A mec complex by MALDI-TOF mass spectrometry. Int J Med Microbiol 2014;304:1018–23. https://doi.org/10.1016/j.ijmm.2014.07.005.

[15] Lau AF, Wang H, Weingarten RA, Drake SK, Suffredini AF, Garfield MK, et al. A rapid matrix-assisted laser desorption ionization-time of flight mass spectrometry-based method for single-plasmid tracking in an outbreak of carbapenem-resistant Enterobacteriaceae. J Clin Microbiol 2014;52:2804–12. https://doi.org/10.1128/JCM.00694-14.

[16] Anhalt JP, Fenselau C. Identification of bacteria using mass spectrometry. Anal Chem 1975;47:219–25. https://doi.org/10.1021/ac60352a007.

[17] Cuénod A, Foucault F, Pflüger V, Egli A. Factors associated with MALDI-TOF mass spectral quality of species identification in clinical routine diagnostics. Front Cell Infect Microbiol 2021;11:646648. https://doi.org/10.3389/fcimb.2021.646648.

[20] Wickham H. ggplot2 [Internet]. New York, NY: Springer New York; 2009. Available at: http://link.springer.com/10.1007/978-0-387-98141-3. [Accessed 29 March 2022].

[21] Mitchell M, Mali S, King CC, Bark SJ. Enhancing MALDI time-of-flight mass spectrometer performance through spectrum averaging. PLoS One 2015;10: e0120932. https://doi.org/10.1371/journal.pone.0120932.

[22] Wittwer M, Lasch P, Drevinek M, Schmoldt S, Indra A, Jacob D, et al. First report: application of MALDI-TOF MS within an external quality assurance exercise for the discrimination of highly pathogenic bacteria from contaminant flora. Appl Biosafety 2012;17:59—63. https://doi.org/10.1177/153567601201700202.

[23] Alatoom AA, Cunningham SA, Ihde SM, Mandrekar J, Patel R. Comparison of direct colony method versus extraction method for identification of gram-positive cocci by use of Bruker Biotyper matrix-assisted laser desorption ionization—time of flight mass spectrometry. J Clin Microbiol 2011;49: 2868—73. https://doi.org/10.1128/JCM.00506-11.

[24] Croxatto A, Prod'hom G, Greub G. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. FEMS Microbiol Rev 2012;36: 380—407. https://doi.org/10.1111/j.1574-6976.2011.00298.x.