frontiers | Research Topics

MALDI-TOF MS in microbiological diagnostics: Future applications beyond identification

Edited by
Karsten Becker and Antonella Lupetti

Published in
Frontiers in Microbiology

ISSN 1664-8714
ISBN 978-2-8325-2374-2
DOI 10.3389/978-2-8325-2374-2

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

MALDI-TOF MS in microbiological diagnostics: Future applications beyond identification

Topic editors

Karsten Becker — University Medicine Greifswald, Germany

Antonella Lupetti — University of Pisa, Italy

Citation

Becker, K., Lupetti, A., eds. (2023). *MALDI-TOF MS in microbiological diagnostics: Future applications beyond identification*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-2374-2

Table of contents

frontiers | Frontiers in Microbiology

TYPE Editorial
PUBLISHED 25 April 2023
DOI 10.3389/fmicb.2023.1204452

OPEN ACCESS

EDITED AND REVIEWED BY
Rustam Aminov,
University of Aberdeen, United Kingdom

*CORRESPONDENCE
Karsten Becker
karsten.becker@med.uni-greifswald.de

RECEIVED 12 April 2023
ACCEPTED 12 April 2023
PUBLISHED 25 April 2023

CITATION
Becker K and Lupetti A (2023) Editorial: MALDI-TOF MS in microbiological diagnostics: future applications beyond identification. *Front. Microbiol.* 14:1204452. doi: 10.3389/fmicb.2023.1204452

Editorial: MALDI-TOF MS in microbiological diagnostics: future applications beyond identification

Karsten Becker[1]* and Antonella Lupetti[2]

[1]Friedrich Loeffler-Institute of Medical Microbiology, University Medicine Greifswald, Greifswald, Germany, [2]Dipartimento di Ricerca Traslazionale e delle Nuove Tecnologie in Medicina e Chirurgia, Università di Pisa, Pisa, Italy

KEYWORDS

antimicrobial resistance, diagnostics, mass spectrometry, MALDI-TOF, susceptibility testing, nano LC-MS/MS

Editorial on the Research Topic
MALDI-TOF MS in microbiological diagnostics: future applications beyond identification

Following a Research Topic on matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) application for susceptibility testing of microorganisms (Becker and Schubert, 2021), three main reasons prompted us to initiate another one on future applications of MALDI-TOF and other MS approaches beyond identification. Firstly, the silent pandemic of multi-drug resistant organisms (MDROs) as designated by the WHO undoubtedly needs more attention. While the development of new antibiotics is receiving due attention, the development of feasible tests for rapid and reliable detection of antimicrobial resistance (AMR) is still passing under the radar of public attention threshold (van Belkum et al., 2013, 2019; Idelevich and Becker, 2019). Beyond the individual treatment, antibiotics could be considered "social" pharmaceuticals as they may also have general effects. Similarly, diagnostic tests used for microbial identification and antibiotic susceptibility testing (AST) deserve to be considered "social" diagnostics with comparable attention and support as antibiotics, since incorrect or late test results can negatively influence both the antibiotic treatment for the patient and MDRO prevention measures, thus leading to avoidable MDRO transmission and increase selection pressure through unnecessarily broad therapy.

The second reason regards sustainability and cost efficiency. MALDI-TOF MS applications beyond identification may ideally expand the application possibilities of a diagnostic device already placed in many routine laboratories (Clark et al., 2013; Schubert and Kostrzewa, 2017; Welker et al., 2019). Not one method alone can meet all of today's requirements for a full, rapid, robust, little personnel-intensive, cost-effective, high-throughput and routine-ready AMR diagnostic approach with excellent sensitivity and specificity. MALDI-TOF MS has the advantage to represent a phenotypic approach, allowing resistance mechanism-independent assays, as known from classical growth-based AST. Besides specific resistance mechanism-based MALDI-TOF MS assays, e.g. addressing enzyme-caused alterations of an antibiotic (Sparbier et al., 2012), universal approaches have been already reported, as the direct-on-target microdroplet growth assay (DOT-MGA) (Idelevich et al., 2018). Further MALDI-TOF MS advantages include rapidity and random access opportunity.

Thirdly, to consider novel MALDI-TOF MS developments not related to identification and AST. The current main efforts to expand the application profile of MALDI-TOF MS beyond identification target its use for AST as reflected by most of the original articles published in this Research Topic. However, they do not refer to the further development and optimization of the technical procedures for AST itself, but to support this process e.g., by machine learning (ML) software tools.

Nowadays, there is an increased effort to take advantage of artificial intelligence (AI) algorithms to analyze the patterns of MALDI-TOF MS peaks for microbial identification and AST (Weis et al., 2020). By the example of the foodborne pathogens *Campylobacter coli* and *C. jejuni*, Feucherolles et al. describe a ML prediction approach as an AMR screening tool for seven antimicrobial resistances. The aim of diagnostic tests is achieving both high sensitivity and specificity. This study shows the potential of this approach, but also the unsolved hurdles to reach the necessary balance between sensitivity and specificity. Here, high sensitivity was chosen as the most important parameter to adjust the threshold score during the tuning part, which led to specificity problems regarding some of the antibiotics tested. The authors concluded that threshold adjustment is vital while elaborating ML pipeline for routine use based on MALDI-TOF mass spectra. Wang H-Y. et al. used ML approaches to construct a prediction model for rapid detection of ciprofloxacin-resistant *Klebsiella pneumoniae* strains based on identified significant biomarkers. However, as with similar ML approaches, they realized that those models cannot be generalized to other microbial species and antibiotics. This limitation has been addressed by Zhang et al. on serial antibiotic resistances prediction. They generated a multi-label prediction model for clindamycin and oxacillin susceptibilities in *Staphylococcus aureus* based on MALDI-TOF MS data. In this context, multi-label learning targets the challenge where each case is represented by a single instance while simultaneously related with a set of labels (Zhang and Zhou, 2014). Wang C. et al. reported a deep learning-based algorithm on a convolutional neural network (CNN) for that a benchmarking study has recently shown that it is able to outperform traditional machine learning methods (Mortier et al., 2021). Here, utilizing the complete information of MALDI-TOF mass spectra for detecting *Enterococcus faecium*, a CNN model to rapidly and accurately predict clinical vancomycin-resistant *E. faecium* (VREfm) was introduced.

Over the past decade, numerous efforts have been made for direct identification and AST of microorganisms in positive blood cultures (BCs). Verroken et al. evaluated the performances of a commercially available system designed to isolate and concentrate microbial cells directly from a positive BC bottle, the so-called "liquid colony," as equivalent of a overnight subcultured colony for identification by MALDI-TOF MS and other purposes.

The polymyxin colistin, a polycationic peptide antimicrobial, is still considered as last line of defense against carbapenemase-producers. Determination of resistance by MS-based assessment of the negatively charged lipopolysaccharide component lipid A as cellular target of the polymyxins is favored by switching the MALDI-TOF-MS device into the negative-ion mode (Dortet et al., 2018). Jeannot et al. evaluated the MALDIxin test for the detection of colistin-resistant *Pseudomonas aeruginosa* clinical strains. All colistin-susceptible and most of the resistant strains were detected in <1 h; the remaining resistant strains were detected after 4-h colistin pre-exposure. However, this procedure needs a spectrometer equipped with a respective switching modality and it creates more effort, e.g., by additional calibration steps, because the positive-ion mode is used for routine diagnostic purposes in clinical microbiology. For this reason, Foglietta et al. developed a MALDI-TOF MS test in positive-ion mode for rapid detection of colistin-resistant *K. pneumoniae* after a 3-h incubation reaching total agreement with the phenotypic reference method.

Gato et al. evaluated the performance of MALDI-TOF MS for rapid detection of carbapenemase activity in *Enterobacterales*, with a standardized procedure with online software for data analysis using carbapenemase-producers and controls in a multicenter study, followed by a 2-month period clinical evaluation. The accuracy ranged from 83.3 to 100% among the eight international centers. Shaidullina et al. evaluated disks containing ertapenem as source of substrate in MALDI-TOF MS-based assay for detection of carbapenemase-producing *Enterobacterales* and found a sensitive, specific and cost-effective alternative. Lu et al. introduced a nanoscale liquid chromatography coupled to tandem MS (nano LC-MS/MS) workflow to detect enterobacterial carbapenemases with a screening method based on peptide groups (14 peptides with 100% specificity, nine peptides with 95–100% sensitivity).

Olson et al. used MALDI-TOF MS to genotype *Moraxella bovis* strains based on two biomarker models to classify strains according to genotype with an overall accuracy of 85.8–100%.

In summary, the main conclusion that can be drawn from this Research Topic is that the development of non-identification MALDI-TOF MS applications is worth continuing. The overriding advantage is its special potential to obtain a lot of information from one single protein spectrum analysis, i.e., species identification, AST results, data of the genetic diversity and, possibly, further information e.g., on virulence factors.

Author contributions

KB and AL managed the Research Topic MALDI-TOF MS in Microbiological Diagnostics: Future Applications Beyond Identification as topic editors. KB and AL wrote the manuscript and both authors reviewed and edited the manuscript.

Funding

The work is funded in part by a grant of the German Federal Ministry of Education and Research (BMBF) (Grant Number 13GW0594B) to KB and in part by the Italian Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) (Grant Number 20177J5Y3P) to AL.

Acknowledgments

The authors gratefully acknowledge helpful discussions with Evgeny A. Idelevich.

Conflict of interest

KB is inventor of a patent application which is owned by the University of Münster and licensed to Bruker Daltonik GmbH.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Becker, K., and Schubert, S. (2021). *MALDI-TOF MS Application for Susceptibility Testing of Microorganisms.* Lausanne: Frontiers Media SA

Clark, A. E., Kaleta, E. J., Arora, A., and Wolk, D. M. (2013). Matrix-assisted laser desorption ionization-time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. *Clin. Microbiol. Rev.* 26, 547–603. doi: 10.1128/CMR.00072-12

Dortet, L., Bonnin, R. A., Pennisi, I., Gauthier, L., Jousset, A. B., Dabos, L., et al. (2018). Rapid detection and discrimination of chromosome- and MCR-plasmid-mediated resistance to polymyxins by MALDI-TOF MS in *Escherichia coli*: the MALDIxin test. *J. Antimicrob. Chemother.* 73, 3359–3367. doi: 10.1093/jac/dky330

Idelevich, E. A., and Becker, K. (2019). How to accelerate antimicrobial susceptibility testing. *Clin. Microbiol. Infect.* 25, 1347–1355. doi: 10.1016/j.cmi.2019.04.025

Idelevich, E. A., Sparbier, K., Kostrzewa, M., and Becker, K. (2018). Rapid detection of antibiotic resistance by MALDI-TOF mass spectrometry using a novel direct-on-target microdroplet growth assay. *Clin. Microbiol. Infect.* 24, 738–743. doi: 10.1016/j.cmi.2017.10.016

Mortier, T., Wieme, A. D., Vandamme, P., and Waegeman, W. (2021). Bacterial species identification using MALDI-TOF mass spectrometry and machine learning techniques: a large-scale benchmarking study. *Comput. Struct. Biotechnol. J.* 19, 6157–6168. doi: 10.1016/j.csbj.2021.11.004

Schubert, S., and Kostrzewa, M. (2017). MALDI-TOF MS in the microbiology laboratory: current trends. *Curr. Issues Mol. Biol.* 23, 17–20. doi: 10.21775/cimb.023.017

Sparbier, K., Schubert, S., Weller, U., Boogen, C., and Kostrzewa, M. (2012). Matrix-assisted laser desorption ionization-time of flight mass spectrometry-based functional assay for rapid detection of resistance against beta-lactam antibiotics. *J. Clin. Microbiol.* 50, 927–937. doi: 10.1128/JCM.05737-11

van Belkum, A., Bachmann, T. T., Lüdke, G., Lisby, J. G., Kahlmeter, G., Mohess, A., et al. (2019). Developmental roadmap for antimicrobial susceptibility testing systems. *Nat. Rev. Microbiol.* 17, 51–62. doi: 10.1038/s41579-018-0098-9

van Belkum, A., Durand, G., Peyret, M., Chatellier, S., Zambardi, G., Schrenzel, J., et al. (2013). Rapid clinical bacteriology and its future impact. *Ann. Lab. Med.* 33, 14–27. doi: 10.3343/alm.2013.33.1.14

Weis, C. V., Jutzeler, C. R., and Borgwardt, K. (2020). Machine learning for microbial identification and antimicrobial susceptibility testing on MALDI-TOF mass spectra: a systematic review. *Clin. Microbiol. Infect.* 20, 2. doi: 10.1016/j.cmi.2020.03.014

Welker, M., Van Belkum, A., Girard, V., Charrier, J. P., and Pincus, D. (2019). An update on the routine application of MALDI-TOF MS in clinical microbiology. *Expert. Rev. Proteomics.* 16, 695–710. doi: 10.1080/14789450.2019.1645603

Zhang, M. L., and Zhou, Z. H. (2014). A Review on Multi-Label Learning Algorithms. *IEEE Trans. Knowl. Data Eng.* 26, 1819–1837. doi: 10.1109/TKDE.2013.39

frontiers in Microbiology

ORIGINAL RESEARCH
published: 31 August 2021
doi: 10.3389/fmicb.2021.725383

Detection of Colistin Resistance in *Pseudomonas aeruginosa* Using the MALDIxin Test on the Routine MALDI Biotyper Sirius Mass Spectrometer

*Katy Jeannot[1,2,3], Katheryn Hagart[4], Laurent Dortet[2,5,6], Markus Kostrzewa[7], Alain Filloux[4], Patrick Plesiat[1,2,3] and Gerald Larrouy-Maumus[4]**

[1]UMR 6249 Chrono-Environnement, UFR Sciences Médicales et Pharmaceutiques, University of Bourgogne-Franche Comté, Besançon, France, [2]French National Reference Centre for Antibiotic Resistance, Besançon, France, [3]Department of Bacteriology, University Hospital of Besançon, Besançon, France, [4]Department of Life Sciences, Faculty of Natural Sciences, MRC Centre for Molecular Bacteriology and Infection, Imperial College London, London, United Kingdom, [5]Department of Bacteriology-Hygiene, Bicêtre Hospital, Assistance Publique – Hôpitaux de Paris, Le Kremlin-Bicêtre, France, [6]EA7361 "Structure, Dynamic, Function and Expression of Broad Spectrum β-lactamases," LabEx Lermit, Faculty of Medecine, Paris-Sud University, Le Kremlin-Bicêtre, France, [7]Bruker Daltonik GmbH, Bremen, Germany

OPEN ACCESS

Edited by:
Antonella Lupetti,
University of Pisa, Italy

Reviewed by:
Dinesh Sriramulu,
Independent Researcher, Chennai, India
Bahman Mirzaei,
Zanjan University of Medical Sciences, Iran

***Correspondence:**
Gerald Larrouy-Maumus
g.larrouy-maumus@imperial.ac.uk

Specialty section:
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

Received: 15 June 2021
Accepted: 05 August 2021
Published: 31 August 2021

Citation:
Jeannot K, Hagart K, Dortet L, Kostrzewa M, Filloux A, Plesiat P and Larrouy-Maumus G (2021) Detection of Colistin Resistance in Pseudomonas aeruginosa Using the MALDIxin Test on the Routine MALDI Biotyper Sirius Mass Spectrometer. Front. Microbiol. 12:725383. doi: 10.3389/fmicb.2021.725383

Colistin is frequently a last resort treatment for *Pseudomonas aeruginosa* infections caused by multidrug-resistant (MDR) and extensively drug resistant (XDR) strains, and detection of colistin resistance is essential for the management of infected patients. Therefore, we evaluated the recently developed MALDIxin test for the detection of colistin resistance in *P. aeruginosa* clinical strains using the routine matrix-assisted laser desorption ionization (MALDI) Biotyper Sirius system. The test is based on the detection by mass spectrometry of modified lipid A by the addition of 4-amino-L-arabinose (L-ara4N) molecules on one or two phosphate groups, in strains resistant to colistin. Overproduction of L-Ara4N molecules is mainly due to the constitutive activation of the histidine kinase (PmrB) or the response regulator (PmrA) following an amino-acid substitution in clinical strains. The performance of the test was determined on a panel of 14 colistin-susceptible and 14 colistin-resistant *P. aeruginosa* clinical strains, the reference strain PAO1 and positive control mutants PmrB (V28G), PmrB (D172), PhoQ (D240–247), and ParR (M59I). In comparison with the broth microdilution (BMD) method, all the susceptible strains ($n = 14$) and 8/14 colistin-resistant strains were detected in less than 1 h, directly on whole bacteria. The remaining resistant strains ($n = 6$) were all detected after a short pre-exposure (4 h) to colistin before sample preparation. Validation of the method on a larger panel of strains will be the next step before its use in diagnostics laboratories. Our data showed that the MALDIxin test offers rapid and efficient detection of colistin resistant *P. aeruginosa* and is thus a valuable diagnostics tool to control the spread of these emerging resistant strains.

Keywords: MALDI mass spectrometry, lipid A, colistin, *Pseudomonas aeruginosa*, clinical isolate

INTRODUCTION

Pseudomonas aeruginosa is an opportunistic pathogen well-known for infections associated with intensive care units. It is one of the most frequent cause of acute pulmonary healthcare-associated infections, and severe infections particularly in immunocompromised patients (Vincent et al., 2009). Its intrinsic resistance to many antibiotics combined with its facility to accumulate a diversity of resistance mechanisms increasingly limits therapeutic options (Horcajada et al., 2019). Thus, polymyxins (polymyxin B or colistin) are used as a last resort for the treatment of *P. aeruginosa* infections caused by multidrug-resistant (MDR) and extensively drug resistant (XDR) strains (Sader et al., 2017; Doi, 2019; Walters et al., 2019; Ahmadian et al., 2020; Mirzaei et al., 2020). Unfortunately, resistance to colistin has emerged. In *P. aeruginosa*, acquired resistance to colistin mostly results from the addition of one or two 4-amino-L-arabinose (L-Ara4N) molecules to the 1 and/or 4' phosphate groups on the lipid A, the anchor of the LPS in the outer membrane (Bhat et al., 1990; Fernandez et al., 2013). While the *P. aeruginosa* genome contains an *eptA* like gene, the addition of phosphoethanolamine to lipid A or the LPS core has not been reported, unlike in *Enterobacterales* and *Acinetobacter baumannii* (Nowicki et al., 2015). The synthesis and the transport of L-Ara4N molecules is encoded by the large *arnBCADTEF-UgD* operon (simplified as *arn*), which is dependent on a complex regulatory network comprising at least 5 two-component systems (PmrA/PmrB, PhoP/PhoQ, ParR/ParS, CprR/CprS, and ColR/ColS; Fernandez et al., 2010, 2012; Needham and Trent, 2013; Nowicki et al., 2015). Furthermore, mutations in chromosomal genes encoding histidine kinase or response regulators of these two-component systems result in constitutive activation of the *arn* operon. However, in *P. aeruginosa* clinical strains, the genetic events most associated with colistin resistance are amino-acid substitutions leading to a gain of function of the PmrB protein (Barrow and Kwon, 2009; Schurek et al., 2009; Bolard et al., 2019). Although, the *mcr* genes have been widely reported in *Enterobacterales*, they have not currently been identified in *P. aeruginosa* strains, except in the chromosome of one clinical isolate strain (*mcr-5*; Snesrud et al., 2018).

Although, strains of *P. aeruginosa* resistant to polymyxins are still rare, their detection is one of the key issues to improve the treatment of patient infected with MDR and XDR *P. aeruginosa* strains (Diekema et al., 2019). Unfortunately, the methods currently available in routine laboratories for the detection of resistance to colistin in *P. aeruginosa*, still rely on bacterial growth in the presence of polymyxins. These procedures require 16–20 h in culture, whether determining susceptibility to colistin minimal inhibitory concentration (MIC) using the broth microdilution method (BMD; reference method), or the colistin broth disk elution and colistin agar test methods recently accepted by the CLSI (2020). Only two test have reported the detection of colistin resistance strains in less than 3 h (Sadek et al., 2020; Sorensen et al., 2020). The first is a biochemical test (Rapid Polymyxin/*Pseudomonas* NP test) based on the change in color of the bromocresol (yellow to purple/violet) following the production of basic metabolites during the growth of the strain in the presence of colistin (Sadek et al., 2020), and a fast lipid analysis technique (FLAT) directly on a matrix-assisted laser desorption ionization (MALDI) plate. However, both approaches have issues: possible misinterpretation of the colorimetric test, and potential cross contamination with the on-target hydrolysis in the FLAT method.

Therefore, there is an urgent need to develop a fast and robust assay to detect colistin-resistant *P. aeruginosa* strains. Recently, we developed a rapid technique using matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) to rapidly detect colistin resistance using whole cells, the MALDIxin test (Dortet et al., 2018a,b). The MALDIxin test has now been optimized for *Escherichia coli*, *Klebsiella pneumoniae*, *A. baumannii*, and *Salmonella enterica* (Dortet et al., 2018a,b, 2019, 2020; Furniss et al., 2019). The aim of this study is to evaluate the performance of the optimized MALDIxin test using a routine MALDI mass spectrometer (in comparison to the BMD method), to detect colistin-resistant *P. aeruginosa*.

MATERIALS AND METHODS

Bacterial Strains

From the *Pseudomonas* collection of the French National Reference Centre for Antibiotic Resistance (Besançon, France), 14 colistin susceptible (MIC ≤2 mg/L) and 14 colistin resistant clinical strains (MIC >2 mg/L) were selected. All the strains were genotypically-unrelated, and isolated from 25 health institutions distributed throughout France. In addition, two *pmrB* mutants (AB8.2, AB16.2), one *parR* mutant (PAOW2), and one *phoQ* mutant (AB8.4) derivate form the *P. aeruginosa* reference strain PAO1 were included as positive controls (**Table 1**; Muller et al., 2011; Bolard et al., 2019). The wild-type strain PAO1 was used as negative control.

Susceptibility Testing

MICs were determined in triplicate by the BMD using colistin sulfate (Sigma Aldrich, Saint Quentin Fallavier, France) and cation-adjusted Mueller Hinton broth (MHB) from Becton Dickinson (Pont-de-Claix, France) according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) recommendations (EUCAST, 2019). Results were interpreted using EUCAST breakpoints (≤2 mg/L; >2 mg/L).

Whole Genome Sequencing

Four clinical strains (185345, 185374, 185819, and 196337) for which the mutation responsible for colistin resistance had not been identified and characterized were genome sequenced. From an overnight bacterial culture, total DNA was extracted using the PureLink Genomic DNA Mini kit (ThermoFischer Scientific). The library preparation using NextEra® XT DNA preparation kit (Illumina, San Diego, CA, United States) and sequencing by Illumina NextSeq500 system (2×150-bp paired end reads) were performed by the "Plateforme de Microbiologie Mutualisée" (PibNet) at Institut Pasteur (Paris, France). Reads were assembled with Shovill-Spades v3.14.0. To identify mutations in genes associated with colistin resistance in *P. aeruginosa*

TABLE 1 | Characteristics and results of the MALDIxin test on *P. aeruginosa* strains.

Name of strain	MIC to colistin (mg/L)	Colistin resistance mechanism	PRR MALDIxin results	References
Reference strains				
PAO1	1	–	0.00 ± 0.00/0.02 ± 0.01*	Tenover, 2000
AB8.2	128	PmrB (V28G)	0.87 ± 0.25	Bolard et al., 2019
AB16.2	128	PmrB (Δ172)	1.17 ± 0.22/2.12 ± 0.07*	Bolard et al., 2019
AB8.4	4	PhoQ (Δ240–247)	0.00 ± 0.00/0.44 ± 0.02*	This study
PAOW2	4	ParR (M59I)	0.29 ± 0.05/0.28 ± 0.01*	Muller et al., 2011
Colistin susceptible clinical strains				
185715	1	–	0.00 ± 0.00/0.00 ± 0.00*	This study
185716	1	–	0.00 ± 0.00/0.00 ± 0.00*	This study
218401	1	–	0.00 ± 0.00/0.00 ± 0.00*	This study
218418	1	–	0.00 ± 0.00/0.00 ± 0.00*	This study
218419	0.5	–	0.00 ± 0.00/0.00 ± 0.00*	This study
218420	1	–	0.00 ± 0.00/0.00 ± 0.00*	This study
218422	0.5	–	0.00 ± 0.00/0.00 ± 0.00*	This study
218423	0.5	–	0.00 ± 0.00/0.00 ± 0.00*	This study
218424	2	–	0.00 ± 0.00/0.07 ± 0.01*	This study
218427	1	–	0.00 ± 0.00/0.00 ± 0.00*	This study
218428	0.5	–	0.00 ± 0.00/0.00 ± 0.00*	This study
218429	0.5	–	0.00 ± 0.00/0.00 ± 0.00*	This study
218435	0.5	–	0.00 ± 0.00/0.00 ± 0.00*	This study
218437	1	–	0.00 ± 0.00/0.03 ± 0.01*	This study
Colistin resistant clinical strains				
142243	128	PmrB, (Q105P)[b]	0.00 ± 0.00/0.22 ± 0.01*	Bolard et al., 2019
152739	16	PmrB (V264A)[a]	0.39 ± 0.04	Bolard et al., 2019
153038	64	PmrB (D47N)[b]	0.67 ± 0.03	Bolard et al., 2019
153091	128	PmrA (L21I)	0.86 ± 0.07	Bolard et al., 2019
163795	128	PmrB (G188D)[b]	0.23 ± 0.01	Bolard et al., 2019
174536	4	PmrB (V136E)[b]	0.00 ± 0.01/2.03 ± 0.13*	Bolard et al., 2019
174660	64	PmrB (G121P, V313A)[b]	0.60 ± 0.44	Bolard et al., 2019
174782	4	PmrB (H33Y)[a]	0.98 ± 0.14	Bolard et al., 2019
175058	4	PmrB (D45N, G362S)[a]	0.24 ± 0.01	Bolard et al., 2019
175101	32	PmrB (R92H, G123S)[b]	0.19 ± 0.07	Bolard et al., 2019
185345	4	–	0.00 ± 0.00/0.87 ± 0.07*	This study
185374	4	–	0.00 ± 0.00/0.22 ± 0.01*	This study
185819	128	PhoQ (R275X)	0.22 ± 0.06	This study
196337	4	–	0.00 ± 0.00/0.34 ± 0.03*	This study

[a]*Mutation Y345H associated with polymorphism in the protein PmrB.*

[b]*Mutations S2P, A4T, V6A, V15I, G68S, and Y345H associated with polymorphism in the protein PmrB.*

–, no mutation has been identified in genes cprS, cprR, parS, parR, pmrA, pmrB, phoP, phoQ, colS, and colR.

**PRR obtained after colistin induction of the strains.*

(*cprR*, *cprS*, *colR*, *colS*, *parR*, *parS*, *phoP*, *phoQ*, *pmrA*, and *pmrB*), sequences from clinical strains were mapped to sequences of reference strains PAO1, and P, A14, and a large collection of 77 clinical strains susceptible to colistin using the NRC bioinformatic pipelines based on SNIPPY v4.6.0.

Nucleotide Sequence Accession Number

The nucleotide sequences reported in this study and corresponding to the entire chromosome of strains 185345, 185374, 185819, and 196337 have been deposited in the GenBank nucleotide database under accession no JAGJAGMWR000 0000000, JAGMWQ0000000000, JAGMWP0000000000, and JAGMWO0000000000, respectively.

MALDIxin Test

Pseudomonas aeruginosa cells were exposed or not for 4-h to a subinhibitory (1/2 MIC) concentration of colistin. A 10 µl inoculation loop of bacteria, grown on Mueller-Hinton agar for 18–24 h, was resuspended in 200 µl of water. Mild-acid hydrolysis was performed on 100 µl of this suspension, by adding 100 µl of 2% *v*/*v* acetic acid and incubating the mixture at 98°C for 30 min. Hydrolyzed cells were centrifuged at 17,000 ×*g* for 2 min, the supernatant was discarded, and the pellet was washed three times with 300 µl of ultrapure water and resuspended to a density of McFarland 20 as measured using a McFarland Tube Densitometer. A volume of 0.4 µl of this suspension was loaded onto the MALDI target plate and immediately overlaid with 1.2 µl of Norharmane matrix (Sigma-Aldrich) solubilized in chloroform/methanol (90:10 *v*/*v*) to a final concentration of 10 mg/ml. For external calibration, 0.5 µl of calibration peptide was loaded along with 0.5 µl of the given calibration matrix (peptide calibration standard II, Bruker Daltonik, Germany). The samples were loaded onto a disposable MSP 96 target polished steel BC (Bruker Part-No. 8280800).

The bacterial suspension and matrix were mixed directly on the target by pipetting then dried gently under a stream of air. The spectra were recorded in the linear negative-ion mode (laser intensity 95%, ion source 1 = 10.00 kV, ion source 2 = 8.98 kV, lens = 3.00 kV, detector voltage = 2,652 V, pulsed ion extraction = 150 ns). Each spectrum corresponded to ion accumulation of 5,000 laser shots randomly distributed on the spot. The spectra obtained were processed with default parameters using FlexAnalysis v.3.4 software (Bruker Daltonik, Germany).

Data Analysis

The negative mass spectrum was scanned between *m*/*z* 1,300 and *m*/*z* 2,000 in the negative linear ion mode. Manual peak picking at masses relevant to colistin resistance was performed on the obtained mass spectra and the corresponding signal intensities at these defined masses was determined. The percentage of modified lipid A was calculated by dividing the sum of the intensities of the lipid A peaks attributed to addition of L-Ara4N (*m*/*z* 1,577.9, *m*/*z* 1,593.9, *m*/*z* 1,708.9, and *m*/*z* 1,724.9) by the intensity of the peaks corresponding to native lipid A (*m*/*z* 1,446.7 and *m*/*z* 1,462.7). All mass spectra were generated and analyzed in technical triplicate (i.e., measurements of each sample were repeated three times).

Statistical Analysis

All experiments were carried out in biological duplicates. Data were compared two-by-two using unpaired Welch's *t*-test. Values of $p < 0.05$ were considered statistically different.

RESULTS AND DISCUSSION

To assess the ability of the MALDIxin test to detect colistin-resistance in *P. aeruginosa* on a MALDI biotyper Sirius system, we tested a panel of 33 *P. aeruginosa* strains, including four isogenic *P. aeruginosa* PAO1 mutants representing the most frequently mutated genes involved in colistin resistance in this species. These included the reference strain PAO1, 14 colistin-resistant clinical strains (MIC from 4 to 128 mg/L), and 14 colistin-susceptible clinical strains (MIC from 0.5 to 2 mg/L; **Table 1**). The mass spectrum of colistin susceptible *P. aeruginosa* reference strain PAO1 is dominated by two peaks assigned to penta-acyl bis-phosphorylated lipid A (**Figure 1A**). The peak at *m/z* 1,446.7 is assigned to penta-acyl bis-phosphorylated lipid A, which corresponds to the presence of one 3-OH-C10:0 fatty acyl chain, three C12:0 fatty acyl chain and one 2-OH C12:0 fatty acyl chain. The lipid A structure at *m/z* 1,462.7 differs from that at *m/z* 1,446.7 only by the addition of one hydroxyl group at 3 position (3-OH) of one C12:0 fatty acyl chain. Both forms have frequently been reported in *P. aeruginosa* strains (Ernst et al., 2006; Moskowitz et al., 2012; **Figure 1D**). In comparison with the parental strain PAO1, four additional peaks at *m/z* 1,577.9, *m/z* 1,593.9, *m/z* 1,708.9, and *m/z* 1,724.9 were observed in the *pmrB* mutant (mutant AB8.2), which is resistant to colistin (**Figure 1B**). The signals at *m/z* 1,577.9, and *m/z* 1,593.9, and at *m/z* 1,708.9 and *m/z* 1,724.9 correspond to the addition of one or two L-Ara4N molecules to the 4′- or/and 1-phosphate groups of the penta-acylated form respectively, resulting in an increase of +131 *m/z* compared to the native lipid A peaks (**Figures 1B,E,F**). Since mutations in genes encoding the two-component systems PmrAB, ParRS, and PhoPQ lead to the addition of L-Ara4N molecules, we did not observe any difference in mass spectra between the *pmrB* (AB8.2 and AB16.2), and *parR* (PAOW2) mutants, compared to the *phoQ* mutant (AB8.4). A similar spectrum is also observed for clinical isolates (**Figures 1C,F**). Based on the MALDIxin profile, we attempted to determine the Polymyxin Resistance Ratio (PRR) of the sum of the intensities of the peaks associated with modified lipid A over the intensity of the peak of native lipid A allowing an accurate distinction between polymyxin-susceptible and polymyxin-resistant isolates. Despite several attempts, no signal corresponding to a modified lipid A was detected. Although, the MIC of colistin was similar in *pmrB* mutants (128 mg/L), the intensity of peaks was clearly higher in *pmrB* mutant AB8.2 (PRR 87 ± 25%) than AB16.2 (PRR 2 ± 0%), indicating that the intensity of peaks is not correlated with resistance level to colistin. The same observation was previously reported for *E. coli*, *Salmonella*, *K. pneumoniae*, and *A. baumannii* (Dortet et al., 2018b, 2019, 2020; Furniss et al., 2019). As expected, the peaks observed in the 14 clinical strains susceptible to colistin did not differ from those obtained with the strain PAO1 (data not shown), and the percentage of modified lipid A was equal to zero (**Table 1**).

Interestingly, in clinical strains resistant to colistin ($n = 14$), the percentage of lipid A modified or PRR ranged from 0 to

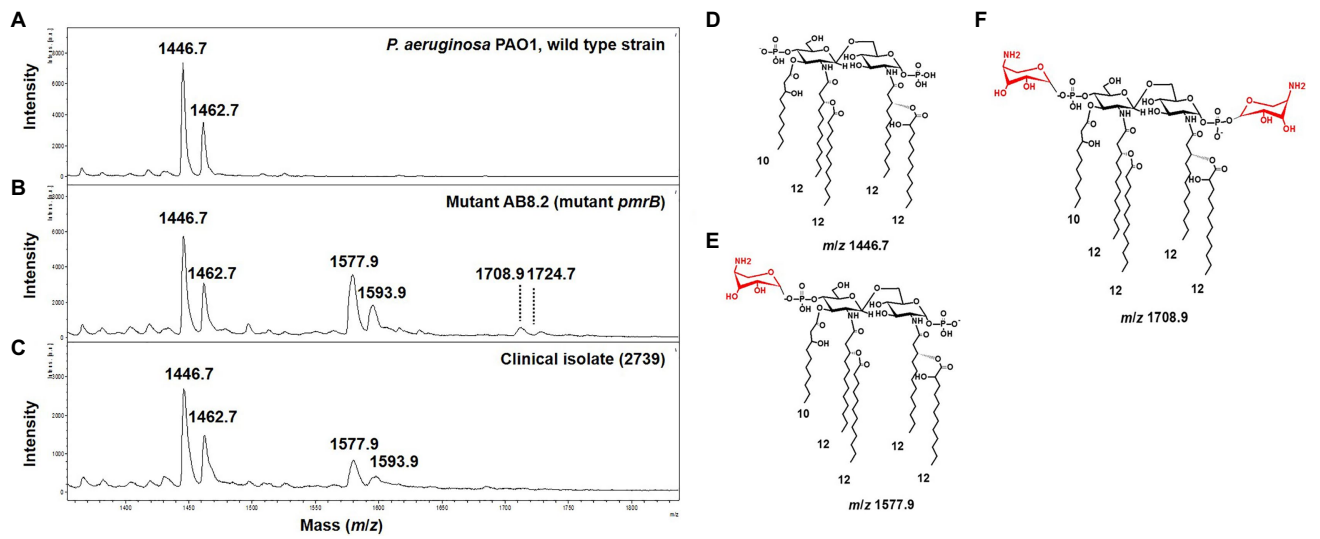


FIGURE 1 | Representative mass spectra of susceptible and modified *Pseudomonas aeruginosa* lipid A acquired using the linear negative-ion mode of a matrix-assisted laser desorption ionization (MALDI) Biotyper Sirius system (Bruker Daltonics). **(A)** Susceptible *P. aeruginosa* PAO1 lipid A is detected as two major peaks at *m/z* 1,446.7 and *m/z* 1,462.7. **(B)** Lipid A from mutant *pmrB* (AB8.2) with additional peaks at *m/z* 1,577.9, *m/z* 1,593.9, *m/z* 1,708.9, and *m/z* 1,724.9 corresponding to 4-amino-L-arabinose (L-Ara4N) addition on the penta-acylated lipid A (peaks at *m/z m/z* 1,446.7 and *m/z* 1,462.7). **(C)** Lipid A from colistin-resistant *P. aeruginosa* clinical isolates is modified by L-Ara4N, which are detected as additional peaks at *m/z* 1,577.9, *m/z* 1,593.9. **(D)** Diagram of *P. aeruginosa* lipid A at *m/z* 1,446.7. **(E)** Diagram of *P. aeruginosa* lipid A at *m/z* 1,577.9. L-Ara4N residue is shown in red. **(F)** Diagram of *P. aeruginosa* lipid A at *m/z* 1,708.9. L-Ara4N residues are shown in red.

90%. Although, five *P. aeruginosa* clinical strains and the *phoQ* mutant (mutant AB8.4) have a MIC higher than the breakpoint for colistin (>2 µg/ml) including one strain with a MIC ≥ 128 mg/L (163795), we did not detect any modifications on the lipid A as reported by their null PRR (**Table 1**). Unlike for MIC determination, bacteria are not exposed to colistin when preparing bacteria for the MALDIxin test. It is likely that induction of the two-component systems PmrAB, CprRS, and ParRS is necessary to detect sufficient modification (beyond background) of lipid A in some strains (Moskowitz et al., 2004; Muller et al., 2011; Fernandez et al., 2012). Therefore, the strain PAO1, the *phoQ* mutant and the six strains were exposed to a sub-inhibitory concentration of colistin (½ MIC) for 4 h, before the MALDIxin test and determination of the percentage of modified lipid A (**Table 1**). While the colistin susceptible reference strain exhibited 2% modified lipid A after colistin exposure, the *phoQ* mutant and colistin resistant clinical strains had more than 20% (**Table 1**; **Figure 2**). All the strains resistant to colistin were detected after a short exposure (4 h) to colistin, confirming that for some strains, the sensitivity of the MALDIxin test can be enhanced by the induction of colistin resistance (**Figure 3**).

Here, we have demonstrated that resistance to colistin can be quickly and easily detected in clinical strains of *P. aeruginosa* using the MALDIxin assay. However, an adaptation of the protocol currently used will be necessary before its use in routine laboratories. Unlike other species such as *K. pneumoniae*, *A. baumannii*, *E. coli*, and *S. enterica*, the signal intensity corresponding to the modified lipid A can be masked in some strains resistant to colistin. The addition of colistin

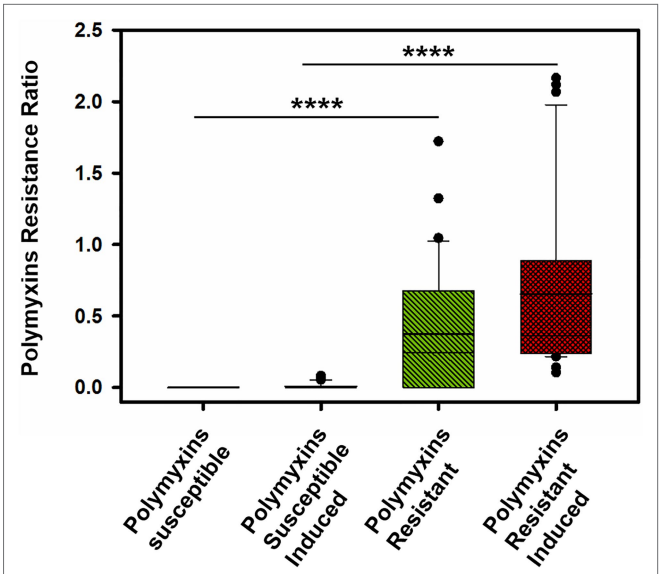


FIGURE 3 | Distribution of the Polymyxin Resistance Ratios (PRRs) for the tested strains before and after induction with colistin for 4-h to a subinhibitory (1/2 MIC; ****$p < 0.0001$).

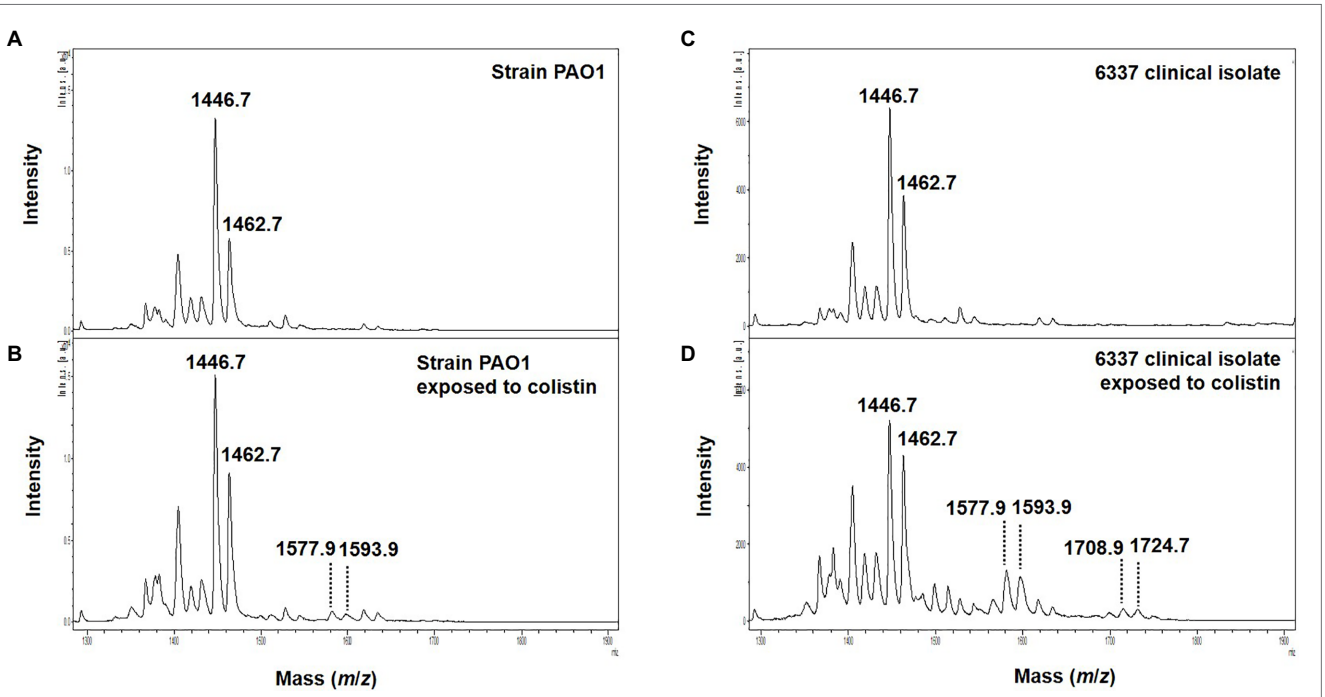


FIGURE 2 | Representative mass spectra of colistin induced susceptible and modified *P. aeruginosa* lipid A acquired using the linear negative-ion mode of a MALDI Biotyper Sirius system (Bruker Daltonics). **(A)** Susceptible *P. aeruginosa* PAO1 lipid A is detected as two major peaks at *m/z* 1,446.7 and *m/z* 1,462.7. **(B)** Colistin induced susceptible *P. aeruginosa* PAO1 lipid A with additional peaks at *m/z* 1,577.9, *m/z* 1,593.9 corresponding to L-Ara4N addition on the penta-acylated lipid A (peaks at *m/z m/z* 1,446.7 and *m/z* 1,462.7). **(C)** Lipid A from colistin-resistant *P. aeruginosa* clinical isolate 6,337 is detected as two major peaks at *m/z* 1,446.7 and *m/z* 1,462.7. **(D)** Colistin induced colistin-resistant *P. aeruginosa* clinical isolate 6,337 lipid A with additional peaks at *m/z* 1,577.9, *m/z* 1,593.9, *m/z* 1,708.9, and *m/z* 1,724.9 corresponding to L-Ara4N addition on the penta-acylated lipid A (peaks at *m/z m/z* 1,446.7 and *m/z* 1,462.7).

during sample preparation phase improves the detection of *P. aeruginosa* strains resistant to colistin. Despite this additional step in sample preparation, the technique remains rapid (less than 5 h), comparing favorably to the BMD for determining the MIC of colistin. The MALDIxin method complements the panel of so-called rapid methods for detecting resistance to colistin in clinical strains of *P. aeruginosa*, including the Rapid Polymyxin/*Pseudomonas* NP test (Sadek et al., 2020). MALDIxin is cost effective since it can be coupled with bacterial identification using the norharmane matrix with the MALDI Biotyper Sirius. One of the limitations of the study resides in the low number of strains tested, and further validation with an expanded panel is required. However, the most frequent mechanism responsible for colistin resistance in *P. aeruginosa* clinical strains (*pmrB* mutants) are included in this study, which supports the use of MALDIxin as a diagnostic for hospitalized patients.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

AUTHOR CONTRIBUTIONS

GL-M, LD, and KJ conceived the study, participated in its design, and performed the experiments. KJ provided the clinical isolates. KJ, KH, LD, MK, AF, PP, and GL-M wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the MRC Confidence in Concept Fund and the ISSF Wellcome Trust Grant 105603/Z/14/Z (GL-M).

ACKNOWLEDGMENTS

We would like to thank Brian Robertson, Imperial College London for careful reading of the manuscript.

REFERENCES

Ahmadian, L., Haghshenas, M. R., Mirzaei, B., Norouzi Bazgir, Z., and Goli, H. R. (2020). Distribution and molecular characterization of resistance gene cassettes containing class 1 integrons in multi-drug resistant (MDR) clinical isolates of *Pseudomonas aeruginosa*. *Infect. Drug Resist.* 13, 2773–2781. doi: 10.2147/IDR.S263759

Barrow, K., and Kwon, D. H. (2009). Alterations in two-component regulatory systems of phoPQ and pmrAB are associated with polymyxin B resistance in clinical isolates of *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 53, 5150–5154. doi: 10.1128/AAC.00893-09

Bhat, R., Marx, A., Galanos, C., and Conrad, R. S. (1990). Structural studies of lipid A from *Pseudomonas aeruginosa* PAO1: occurrence of 4-amino-4-deoxyarabinose. *J. Bacteriol.* 172, 6631–6636. doi: 10.1128/jb.172.12.6631-6636.1990

Bolard, A., Schniederjans, M., Haussler, S., Triponney, P., Valot, B., Plesiat, P., et al. (2019). Production of norspermidine contributes to aminoglycoside resistance in pmrAB mutants of *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 63:e01044-19. doi: 10.1128/AAC.01044-19

Diekema, D. J., Hsueh, P. R., Mendes, R. E., Pfaller, M. A., Rolston, K. V., Sader, H. S., et al. (2019). The microbiology of bloodstream infection: 20-year trends from the SENTRY antimicrobial surveillance program. *Antimicrob. Agents Chemother.* 63:e00355-19. doi: 10.1128/AAC.00355-19

Doi, Y. (2019). Treatment options for carbapenem-resistant gram-negative bacterial infections. *Clin. Infect. Dis.* 69(Suppl. 7), S565–S575. doi: 10.1093/cid/ciz830

Dortet, L., Bonnin, R. A., Le Hello, S., Fabre, L., Bonnet, R., Kostrzewa, M., et al. (2020). Detection of colistin resistance in *Salmonella enterica* using MALDIxin test on the routine MALDI Biotyper Sirius mass spectrometer. *Front. Microbiol.* 11:1141. doi: 10.3389/fmicb.2020.01141

Dortet, L., Bonnin, R. A., Pennisi, I., Gauthier, L., Jousset, A. B., Dabos, L., et al. (2018a). Rapid detection and discrimination of chromosome- and MCR-plasmid-mediated resistance to polymyxins by MALDI-TOF MS in *Escherichia coli*: the MALDIxin test. *J. Antimicrob. Chemother.* 73, 3359–3367. doi: 10.1093/jac/dky330

Dortet, L., Broda, A., Bernabeu, S., Glupczynski, Y., Bogaerts, P., Bonnin, R., et al. (2019). Optimization of the MALDIxin test for the rapid identification of colistin resistance in *Klebsiella pneumoniae* using MALDI-TOF MS. *J. Antimicrob. Chemother.* 75, 110–116. doi: 10.1093/jac/dkz405

Dortet, L., Potron, A., Bonnin, R. A., Plesiat, P., Naas, T., Filloux, A., et al. (2018b). Rapid detection of colistin resistance in *Acinetobacter baumannii* using MALDI-TOF-based lipidomics on intact bacteria. *Sci. Rep.* 8:16910. doi: 10.1038/s41598-018-35041-y

Ernst, R. K., Adams, K. N., Moskowitz, S. M., Kraig, G. M., Kawasaki, K., Stead, C. M., et al. (2006). The *Pseudomonas aeruginosa* lipid A deacylase: selection for expression and loss within the cystic fibrosis airway. *J. Bacteriol.* 188, 191–201. doi: 10.1128/JB.188.1.191-201.2006

EUCAST (2019). The European Committee on antimicrobial susceptibility testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 9.0, 2019. Available at: https://eucast.org/clinical_breakpoints/

Fernandez, L., Alvarez-Ortega, C., Wiegand, I., Olivares, J., Kocincova, D., Lam, J. S., et al. (2013). Characterization of the polymyxin B resistome of *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 57, 110–119. doi: 10.1128/AAC.01583-12

Fernandez, L., Gooderham, W. J., Bains, M., McPhee, J. B., Wiegand, I., and Hancock, R. E. (2010). Adaptive resistance to the "last hope" antibiotics polymyxin B and colistin in *Pseudomonas aeruginosa* is mediated by the novel two-component regulatory system ParR-ParS. *Antimicrob. Agents Chemother.* 54, 3372–3382. doi: 10.1128/AAC.00242-10

Fernandez, L., Jenssen, H., Bains, M., Wiegand, I., Gooderham, W. J., and Hancock, R. E. (2012). The two-component system CprRS senses cationic peptides and triggers adaptive resistance in *Pseudomonas aeruginosa* independently of ParRS. *Antimicrob. Agents Chemother.* 56, 6212–6222. doi: 10.1128/AAC.01530-12

Furniss, R. C. D., Dortet, L., Bolland, W., Drews, O., Sparbier, K., Bonnin, R. A., et al. (2019). Detection of colistin resistance in *Escherichia coli* by use of the MALDI Biotyper Sirius mass spectrometry system. *J. Clin. Microbiol.* 57:e01427-19. doi: 10.1128/JCM.01427-19

Horcajada, J. P., Montero, M., Oliver, A., Sorli, L., Luque, S., Gomez-Zorrilla, S., et al. (2019). Epidemiology and treatment of multidrug-resistant and extensively drug-resistant *Pseudomonas aeruginosa* infections. *Clin. Microbiol. Rev.* 32:e00031-19. doi: 10.1128/CMR.00031-19

Mirzaei, B., Bazgir, Z. N., Goli, H. R., Iranpour, F., Mohammadi, F., and Babaei, R. (2020). Prevalence of multi-drug resistant (MDR) and extensively drug-resistant (XDR) phenotypes of *Pseudomonas aeruginosa* and *Acinetobacter baumannii* isolated in clinical samples from northeast of Iran. *BMC Res. Notes* 13:380. doi: 10.1186/s13104-020-05224-w

Moskowitz, S. M., Brannon, M. K., Dasgupta, N., Pier, M., Sgambati, N., Miller, A. K., et al. (2012). PmrB mutations promote polymyxin resistance of *Pseudomonas aeruginosa* isolated from colistin-treated cystic fibrosis patients. *Antimicrob. Agents Chemother.* 56, 1019–1030. doi: 10.1128/AAC.05829-11

Moskowitz, S. M., Ernst, R. K., and Miller, S. I. (2004). PmrAB, a two-component regulatory system of *Pseudomonas aeruginosa* that modulates resistance to cationic antimicrobial peptides and addition of aminoarabinose to lipid A. *J. Bacteriol.* 186, 575–579. doi: 10.1128/JB.186.2.575-579.2004

Muller, C., Plesiat, P., and Jeannot, K. (2011). A two-component regulatory system interconnects resistance to polymyxins, aminoglycosides, fluoroquinolones, and beta-lactams in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 55, 1211–1221. doi: 10.1128/AAC.01252-10

Needham, B. D., and Trent, M. S. (2013). Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis. *Nat. Rev. Microbiol.* 11, 467–481. doi: 10.1038/nrmicro3047

Nowicki, E. M., O'Brien, J. P., Brodbelt, J. S., and Trent, M. S. (2015). Extracellular zinc induces phosphoethanolamine addition to *Pseudomonas aeruginosa* lipid A via the ColRS two-component system. *Mol. Microbiol.* 97, 166–178. doi: 10.1111/mmi.13018

Sadek, M., Tinguely, C., Poirel, L., and Nordmann, P. (2020). Rapid Polymyxin/*pseudomonas* NP test for rapid detection of polymyxin susceptibility/resistance in *Pseudomonas aeruginosa*. *Eur. J. Clin. Microbiol. Infect. Dis.* 39, 1657–1662. doi: 10.1007/s10096-020-03884-x

Sader, H. S., Huband, M. D., Castanheira, M., and Flamm, R. K. (2017). *Pseudomonas aeruginosa* antimicrobial susceptibility results from four years (2012 to 2015) of the international network for optimal resistance monitoring program in the United States. *Antimicrob. Agents Chemother.* 61:e02252-16. doi: 10.1128/AAC.02252-16

Schurek, K. N., Sampaio, J. L., Kiffer, C. R., Sinto, S., Mendes, C. M., and Hancock, R. E. (2009). Involvement of pmrAB and phoPQ in polymyxin B adaptation and inducible resistance in non-cystic fibrosis clinical isolates of *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 53, 4345–4351. doi: 10.1128/AAC.01267-08

Snesrud, E., Maybank, R., Kwak, Y. I., Jones, A. R., Hinkle, M. K., and McGann, P. (2018). Chromosomally encoded mcr-5 in colistin-nonsusceptible *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 62:e00679-18. doi: 10.1128/AAC.00679-18

Sorensen, M., Chandler, C. E., Gardner, F. M., Ramadan, S., Khot, P. D., Leung, L. M., et al. (2020). Rapid microbial identification and colistin resistance detection via MALDI-TOF MS using a novel on-target extraction of membrane lipids. *Sci. Rep.* 10:21536. doi: 10.1038/s41598-020-78401-3

Vincent, J. L., Rello, J., Marshall, J., Silva, E., Anzueto, A., Martin, C. D., et al. (2009). International study of the prevalence and outcomes of infection in intensive care units. *JAMA* 302, 2323–2329. doi: 10.1001/jama.2009.1754

Walters, M. S., Grass, J. E., Bulens, S. N., Hancock, E. B., Phipps, E. C., Muleta, D., et al. (2019). Carbapenem-resistant *Pseudomonas aeruginosa* at US emerging infections program sites, 2015. *Emerg. Infect. Dis.* 25, 1281–1288. doi: 10.3201/eid2507.181200

Conflict of Interest: LD, AF, and GL-M are co-inventors of the MALDIxin test for which a patent has been filed by Imperial Innovations (WO2018158573). MK is employee of Bruker, the manufacturer of the MALDI-TOF MS used in this study.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 27 January 2022
doi: 10.3389/fmicb.2021.789731

Multicenter Performance Evaluation of MALDI-TOF MS for Rapid Detection of Carbapenemase Activity in Enterobacterales: The Future of Networking Data Analysis With Online Software

Eva Gato[1†], Ahalieyah Anantharajah[2†], Manuel J. Arroyo[3†], María José Artacho[4†], Juan de Dios Caballero[5†], Ana Candela[6†], Kateřina Chudějová[7†], Ignacio Pedro Constanso[8†], Cristina Elías[4,9†], Javier Fernández[10†], Jesús Jiménez[3†], Pilar Lumbreras[10†], Gema Méndez[3†], Xavier Mulet[11†], Patricia Pérez-Palacios[12†], Belén Rodríguez-Sánchez[6†], Rafael Cantón[5], Jaroslav Hrabák[7], Luis Mancera[3], Luis Martínez-Martínez[4,9,13], Antonio Oliver[11], Álvaro Pascual[12], Alexia Verroken[2], Germán Bou[1‡] and Marina Oviaño[1‡]*

[1] Servicio de Microbiología, Red Española de Investigación en Patología Infecciosa, Instituto de Investigación Biomédica da Coruña, CIBER de Enfermedades Infecciosas (CIBERIFEC), Instituto de Salud Carlos III, Complejo Hospitalario Universitario A Coruña, A Coruña, Spain, [2] Service de Microbiologie, Cliniques Universitaires Saint-Luc, Brussels, Belgium, [3] Clover BioSoft, Granada, Spain, [4] Unidad de Gestión Clínica de Microbiología, Red Española de Investigación en Patología Infecciosa, Hospital Universitario Reina Sofía, CIBER de Enfermedades Infecciosas (CIBERIFEC), Instituto de Salud Carlos III, Córdoba, Spain, [5] Servicio de Microbiología, Red Española de Investigación en Patología Infecciosa, Hospital Universitario Ramón y Cajal, Instituto Ramón y Cajal de Investigación Sanitaria, CIBER de Enfermedades Infecciosas (CIBERIFEC), Instituto de Salud Carlos III, Madrid, Spain, [6] Servicio de Microbiología, Hospital General Universitario Gregorio Marañón, Madrid, Spain, [7] Department of Microbiology, Faculty of Medicine, University Hospital in Pilsen, Charles University, Pilsen, Czechia, [8] Servicio de Análisis Clínicos, Complejo Hospitalario Universitario A Coruña, A Coruña, Spain, [9] Instituto Maimónides de Investigación Biomédica de Córdoba, Córdoba, Spain, [10] Servicio de Microbiología, Hospital Universitario Central de Asturias, Instituto de Investigación Sanitaria del Principado de Asturias, Oviedo, Spain, [11] Servicio de Microbiología, Hospital Universitario Son Espases, Red Española de Investigación en Patología Infecciosa, CIBER de Enfermedades Infecciosas (CIBERIFEC), Instituto de Salud Carlos III, Palma, Spain, [12] Unidad Clínica de Enfermedades Infecciosas y Microbiología Clínica, CSIC, Red Española de Investigación en Patología Infecciosa, Hospital Universitario Virgen Macarena, Instituto de Biomedicina de Sevilla, CIBER de Enfermedades Infecciosas (CIBERIFEC), Instituto de Salud Carlos III, Universidad de Sevilla, Seville, Spain, [13] Departamento de Química Agrícola, Edafología y Microbiología, Universidad de Córdoba, Córdoba, Spain

In this study, we evaluate the performance of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) for rapid detection of carbapenemase activity in Enterobacterales in clinical microbiology laboratories during a multicenter networking validation study. The study was divided into three different stages: "software design," "intercenter evaluation," and "clinical validation." First, a standardized procedure with an online software for data analysis was designed. Carbapenem resistance was detected by measuring imipenem hydrolysis and the results were automatically interpreted using the Clover MS data analysis software (Clover BioSoft, Spain). Second, a series of 74 genotypically characterized Enterobacterales (46 carbapenemase-producers and 28 non carbapenemase-producers) were analyzed in 8 international centers to ensure the reproducibility of the method. Finally, the methodology was evaluated independently in all centers during a 2-month period and results were compared with the reference standard for carbapenemase detection

OPEN ACCESS

Edited by:
Antonella Lupetti, University of Pisa, Italy

Reviewed by:
Miriam Cordovana, Bruker Daltonik GmbH, Germany
Dortet Laurent, Bicêtre Hospital, France

***Correspondence:**
Marina Oviaño
marina.oviano.garcia@sergas.es

† These authors have contributed equally to this work and share first authorship
‡ These authors have contributed equally to this work and share last authorship

Specialty section:
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

Received: *05 October 2021*
Accepted: *10 December 2021*
Published: *27 January 2022*

Citation:
Gato E, Anantharajah A, Arroyo MJ, Artacho MJ, Caballero JdD, Candela A, Chudějová K, Constanso IP, Elías C, Fernández J, Jiménez J, Lumbreras P, Méndez G, Mulet X, Pérez-Palacios P, Rodríguez-Sánchez B, Cantón R, Hrabák J, Mancera L, Martínez-Martínez L, Oliver A, Pascual Á, Verroken A, Bou G and Oviaño M (2022) Multicenter Performance Evaluation of MALDI-TOF MS for Rapid Detection of Carbapenemase Activity in Enterobacterales: The Future of Networking Data Analysis With Online Software. Front. Microbiol. 12:789731. doi: 10.3389/fmicb.2021.789731

used in each center. The overall agreement rate relative to the reference method for carbapenemase resistance detection in clinical samples was 92.5%. The sensitivity was 93.9% and the specificity, 100%. Results were obtained within 60 min and accuracy ranged from 83.3 to 100% among the different centers. Further, our results demonstrate that MALDI-TOF MS is an outstanding tool for rapid detection of carbapenemase activity in Enterobacterales in clinical microbiology laboratories. The use of a simple in-house procedure with online software allows routine screening of carbapenemases in diagnostics, thereby facilitating early and appropriate antimicrobial therapy.

Keywords: MALDI-TOF MS, carbapenemases enzymes, resistance detection, imipenem, clinical microbiology

INTRODUCTION

Carbapenemase-producing Enterobacterales (CPE) have been recognized by a number of national and international health organizations to represent a major threat to global health (Rodríguez-Baño et al., 2018; U.S. Centers for Disease Control and Prevention, 2019; European Center for Disease Prevention and Control, 2020). Therefore, the search for new, rapid and effective tools for detecting these bacteria in microbiology laboratories is essential to enable early targeted antibiotic therapy. Numerous phenotypic, lateral flow and DNA-based methods have been used in the laboratory with the aim of detecting CPE (Hrabák et al., 2014). The results of culture-based methods are generally available within 24 h of isolation of the bacteria from clinical samples, whereas more expensive genotypic test results are available within hours. Moreover, DNA-based methods can only detect a predefined range of carbapenem-encoding genes, the presence of which does not guarantee expression (Decousser et al., 2015; Dortet and Naas, 2017; Oueslati et al., 2019).

Since the introduction of MALDI-TOF in clinical microbiology laboratories, approaches to detect carbapenemases have been made by using MALDI-TOF MS-based measurement of the hydrolysis of different antibiotics (Clark et al., 2013; Papagiannitsis et al., 2015; Oviaño and Bou, 2018; Oviaño et al., 2019). However, several in-house MALDI-TOF MS assays that use different antibiotic combinations, buffers and variable incubation times, ranging from 15 min to 4 h (Burckhardt and Zimmermann, 2011; Hrabák et al., 2011; Lasserre et al., 2015; Monteferrante et al., 2016; Carvalhaes et al., 2018), have been described. In its guidelines for detection of resistance mechanisms, version 2.0 of July 2017, the EUCAST committee commented on the detection of carbapenem hydrolysis by MALDI-TOF MS. Although the method was recommended for detection of carbapenemases in Enterobacterales, the document highlighted the lack of evaluation of the procedure in a multicenter study design or in studies involving multiple individual centers as a limitation to its use. Hence, standardization of the methodology is required before universal application of MALDI-TOF for the detection of carbapenemase activity based on hydrolysis of a carbapenem antibiotic.

In line with the standardization and automation of the process, the MBT STAR®-Carba IVD kit (Bruker Daltonik, Germany) and associated MBT STAR®-BL IVD software have become available for a fully developed workflow, with a benchmark carbapenem antibiotic, in Bruker equipment (Anantharajah et al., 2019; Cordovana et al., 2020). However, as a result of the increasing need to share data among different laboratories we have developed an online software for mass spectra data analysis that is universally compatible with any mass spectra format derived from any MALDI-TOF commercial brand, so it is independent from suppliers other than the commercial companies who distribute MALDI-TOF MS equipment.

The aim of this study was to conduct a multicenter evaluation of the accuracy and applicability of MALDI-TOF MS for rapid detection of CPE based on the hydrolysis of imipenem with a standardized in-house procedure and subsequent automated interpretation of the results with online mass spectra data analysis software.

MATERIALS AND METHODS

Study Design

The study was divided into three different stages. The first part, i.e., "software design," was carried out to implement online mass spectra data analysis software. To this end, the imipenem hydrolysis assay conditions were the same as those established by Oviaño et al. (2019), using a collection of 74 genotypically characterized CPE and data analysis with the Clover MS data analysis software (Clover BioSoft, Spain)[1]. The results were compared with those obtained with the MBT STAR®-Carba IVD Kit (Bruker Daltonik) and STAR®-BL IVD software. The principle of the assay is that imipenem is inactivated due to hydrolysis of the β-lactam ring by bacteria expressing carbapenemase-hydrolyzing enzymes. The hydrolysis reaction modifies the structure of the antibiotic, so disappearance in the native mass peaks (300 and 489 m/z) can be detected by MALDI-TOF. Thus, a positive hydrolysis reaction denotes the presence of a carbapenemase in bacteria. This first part of the study process was accomplished by the Complejo Hospitalario Universitario A Coruña (AC), A Coruña, considered the reference center.

The second part of the study, i.e., "intercenter evaluation," was designed to standardize the methodology and evaluate the reproducibility among the different centers participating in the survey with the proposed workflow. The same isolates used in

[1]https://platform.clovermsdataanalysis.com/login

the first part of the study were analyzed in the eight participating centers: the Hospital General Universitario Gregorio Marañón (GM), Madrid; the Hospital Universitario Ramón y Cajal (RC), Madrid; the Hospital Universitario Reina Sofía (RS), Córdoba; the Hospital Universitario Virgen Macarena (VM), Seville; the Hospital Universitario Son Espases (SE), Palma; the Hospital Universitario Central de Asturias (CA), Oviedo; the University Hospital Plzeň (PZ), Plzeň; and the Cliniques Universitaires Saint-Luc (SL), Brussels. All reagents were provided by the reference center and solutions were prepared in all centers according to instructions. So, the centers were given a precise protocol with the standardized procedure and an educational video for software management (**Supplementary Data Sheet 1** and **Supplementary Video 1**). Technical support regarding the laboratory procedure and general attention together with software management were available during the study. The results of the MALDI-TOF assay were compared with those obtained by molecular characterization of the isolates.

The last part of the study, i.e., "clinical validation," was designed to evaluate the proposed methodology and data analysis in real clinical settings. During a 2 months period, all isolates suspected by EUCAST screening criteria (i.e., with a meropenem or ertapenem MIC > 0.125 mg/L) to be carrying a carbapenemase enzyme (maximum 20) were evaluated in parallel in the participating centers using the MALDI-TOF MS assay based on imipenem hydrolysis. The results obtained were compared with those yielded by the reference method used in each clinical laboratory for CPE detection.

Bacterial Isolates

A representative collection of 76 non-duplicated clinical isolates (retrospectively obtained from our own collection of isolates belonging to AC) was used in the first and second parts of the study. These included 74 isolates (46 CPE and 28 non-CPE) and 2 control strains. The control strains used were a *E. coli* ATCC 25922 (negative control) and a PCR-confirmed VIM-producing *Citrobacter freundii* (positive control). The 74 test species included in the study were 39 *Klebsiella pneumoniae,* 26 *Escherichia coli*, 8 *Enterobacter cloacae* complex, and 1 *Klebsiella oxytoca*. The isolates were characterized by PCR and sequencing (Oteo et al., 2013, 2016; Dortet et al., 2016; Oviaño et al., 2017). The collection included 46 CPE isolates, being 10 bla_{KPC}, 3 bla_{IMP}, 5 bla_{NDM}, 10 bla_{VIM} and 18 $bla_{OXA-48-like}$, and 28 non-CPE isolates: 1 bla_{CIT}, 1 bla_{SHV}, 2 bla_{CMY}, 3 bla_{FOX}, 1 bla_{K-1}, 14 bla_{CTX-M}, and 6 fully susceptible isolates (see **Supplementary Table 1** for further details). All non-CPE isolates have carbapenems MIC below the cut-off for screening according to EUCAST criteria, excepting two isolates, one *E. coli* and 1 *K. pneumoniae*. Both were tested by PCR and sequencing and no carbapenemase enzyme was found.

In the final part of the study, each center prospectively evaluated (following EUCAST screening guidelines) suspected CPE isolates that emerged in routine testing in clinical microbiology laboratories, i.e., the number of isolates that emerged during the 2 months period or up to 20 isolates. One isolate per patient included. The isolates were derived from different clinical samples. Results were compared with those obtained with the reference method used in each laboratory.

MALDI-TOF MS Acquisition

Bacterial isolates were stored at −80°C, in a small vial containing glass cryopearls (Deltalab, Barcelona, Spain). On day 1, isolates were thawed on a blood agar plate by removing one of the pearls from the tube with a sterile loop and rolling it on the plate. After incubation for 18 h, the isolates were subcultured for another 18 h on a blood agar plate for analysis under standard conditions. The samples were incubated in an aerobic atmosphere at 37°C. Briefly, bacteria filling a 1 μl inoculation loop were suspended in 50 μl of solution (10 mM NH_4CO_3, 10 μg/ml $ZnCl_2$, 0.001% SDS; pH 8) containing 0.5 mg/mL of imipenem (Sigma-Aldrich, Germany) and incubated for 30 min at 37°C with slow agitation (300 rpm) (Oviaño et al., 2019). The samples were then centrifuged at 14,000 rpm, and the supernatant was applied to a MALDI-TOF MS target plate. Each sample was spotted on the plate in duplicate. Once dried, 1 μl of matrix [Matrix IVD HCCA-portioned (Bruker Daltonik) spiked with 1 ppm/μl of reserpine (Sigma-Aldrich, Germany)] was applied to each spot. All runs were performed in the presence of the positive and negative controls treated in the same way as the samples. Because of the instability of imipenem, the antibiotic solution must be prepared just before use, or freezed at −20°C for 1 week or at −70°C for a month. The negative control is useful to evaluate spontaneous hydrolysis of the antibiotic.

The configuration of the equipment was the Research Use Only version as the IVD version does not allow access to the modification of parameters in the FlexControl software. Appropriate calibration was conducted before each run using a mixture of bradykinin [1-5] and [1-7] at 35 μM (Sigma-Aldrich). The mixture was spotted and once dried 1 μl of matrix was added on top of each spot. The spectra were acquired after the mixture was dried.

The mass spectra were obtained using a MALDI Biotyper® Smart (Bruker Daltonik) system, with Flex Control 3.4 software. Three methods were created for spectral acquisition: MBT_ATB.par, MBT_ATB_AutoX.axe and MBT_ATB.mcl. The operational mass range was between 100 and 1,000 m/z in the linear positive mode. The mass peaks were acquired in 40 shot steps to produce 240 satisfactory shots, and the resolution of the mass peaks selected was higher than 300. The movement of the laser in the spot followed a large spiral.

The parameters considered in the calibration are listed in **Table 1**.

MALDI-TOF MS Data Analysis

In the first part of the study, i.e., "software design," the results were analyzed in parallel, with our in-house imipenem hydrolysis procedure and Clover MS data analysis software (Clover BioSoft, Spain) and compared with those obtained with the MBT STAR®-Carba IVD Kit (Bruker Daltonik) and the STAR®-BL IVD software.

Instructions for using Clover MS data analysis software (Clover BioSoft, Spain) in the Carbapenem Hydrolysis Detection Analysis module are provided in **Supplementary Data Sheet 1**

and **Supplementary Video 1**. The software automatically performs the baseline subtraction and subsequently calculates the ratio of hydrolysis (RH) of imipenem, based on the ratios of the mass peaks of imipenem (300 m/z) and imipenem complexed with the matrix (489 m/z) and the internal standard, i.e., reserpine (607 m/z) (**Figure 1**). The mass peak at 300 m/z is usually less intense than the 489 m/z as imipenem tends to stabilize forming an adduct with the matriz (Oviaño and Bou, 2018). The results were normalized according to the controls. Results are displayed in a .pdf report (**Supplementary Data Sheet 2**).

RESULTS

Software Design

In the STAR®-BL IVD software, (a) the cut-off values used were those recommended by the manufacturer. In the Clover MS data analysis software (Clover BioSoft), the RH was calculated in two different ways: (b) by using the intensity of the imipenem mass peaks and (c) by using the area under the curve (AUC) of the peaks. A receiver operating characteristic (ROC) curve was calculated for the three methods, yielding AUC values of (a) 0.977 (0.939–1), (b) 0.994 (0.985–1), and (c) 0.990 (0.976–1), with a 95% confidence interval. As the best results were yielded by the analysis that integrated the mass peaks using intensity (b), the spectra were then treated in this way in the rest of the study. The cut-off values for the in-house developed assay were determined using the ROC curve and the Youden index. The cut-off value established for positivity in the imipenem hydrolysis assay was RH values equal to or above 0.5. Values below 0.2 were considered to indicate negative hydrolysis. A gray or intermediate zone was established between 0.2 and 0.5, which required further testing, including prolonged incubation times, or confirmation by other techniques. The negative and positive controls were each analyzed 20 times in the same run to ensure repeatability. A minimum delta RH value between the positive and negative control of three times the value of the standard deviation was established. The value was established in 0.6. Thus, the positive control should have values $\geq$0.5, the negative control values $<$0.2 and the minimum delta RH value should be 0.6.

For comparing which technique yielded best results and if these differences were statistically significant we analyzed the ROC curve. As the AUC for the STAR®-BL IVD software (a) was 0.977 and this value is out of the 95% confidence interval of the AUC of our developed software (b), (0.985–1) we can conclude that the performance of our developed Clover MS data analysis software (Clover BioSoft) is significantly better when using the intensity of the mass peaks.

Continuous variables are presented as mean (95% confidence interval). According to the cut-off values recommended by the manufacturers, 95.06% of the isolates were correctly classified in (a) the STAR®-BL IVD software, with 92.00% (79.89–97.41) sensitivity and 100% (86.27–99.71) specificity, whereas using (b) the Clover MS data analysis software (Clover BioSoft), 96.30% of the isolates were correctly classified, with 94.00% (82.46–98.44) sensitivity and 100% (86.27–99.71) specificity. The correlation between the two different procedures and data analysis was calculated with the Spearman correlation coefficient, yielding a ρ value of 0.971.

TABLE 1 | MALDI-TOF MS calibration parameters.

Name	m/z	Resolution	Intensity
$[HCCA+H]^+$	190.050	>300	>3000
$[2HCCA+H]^+$	379.092	>500	>3000
$[Bradykinin(1-5)]^+$	573.314	>700	>2000
$[Reserpine]^+$	607.680	>800	>3000
$[Bradykinin(1-7)]^+$	757.400	>1000	>3000

The mass peaks considered for calibration, including the exact masses, resolutions and intensities.

Intercenter Evaluation

In the second part of the study, the same isolates used in the first part were analyzed by the eight international centers participating in the study (**Supplementary Table 2**). Two isolates, i.e., OXA-244-producing *K. pneumoniae* and OXA-232-producing *E. coli* although being classified as CPE according to literature, both enzymes have a low carbapenemase activity, even lower than OXA-48 enzyme (Potron et al., 2013; Hoyos-Mallecot et al., 2017; Pitout et al., 2019). Therefore, data on these isolates were not included in the statistical, as were considered outliers. However, the results were consistent with a low level of enzymatic activity. Among the nine centers, one center provided an intermediate result for the OXA-244-producing *K. pneumoniae*, and three centers provided a negative hydrolysis result. In the case of the OXA-232-producing *E. coli*, one center provided an intermediate result and five centers provided negative results for imipenem hydrolysis.

Once all of the results were obtained, the ROC curve was recalculated with the data obtained in our center (CHUAC) plus the 8 participating centers, to strengthen the statistical power of the analysis. For statistical calculations, 1296 spectra were analyzed from duplicate measurements of the 72 isolates included in the study and measured in the nine centers. The cut-off value for positivity was finally established for RH values $\geq$0.4; negative hydrolysis values yielded RH $<$ 0.2, and the gray zone, with intermediate values, was established for values of RH between 0.2 and 0.4. Thus, only the positivity cut-off changed slightly. The value of the positive control was adjusted to the new value. The overall sensitivity of the procedure was 95.5% (92.9–97.1), with 100% (98.5–100) specificity. The positive predictive value was 100% (99.0–100) and the negative value, 93.3% (89.8–95.7).

For statistical calculations, full agreement was considered when the MALDI-TOF assay provided an imipenem positive hydrolysis result for CPE or a negative hydrolysis result for non-CPE isolates. Three types of errors were also considered. Errors were considered minor when the result of the MALDI-TOF assay was in the intermediate category and the isolate was either CPE or non-CPE. Errors were considered major when the result of the MALDI-TOF assay was positive for imipenem hydrolysis for a non-CPE isolate and very major errors were considered when a negative hydrolysis result was obtained by MALDI-TOF for a CPE. The overall agreement rate was 94.9%

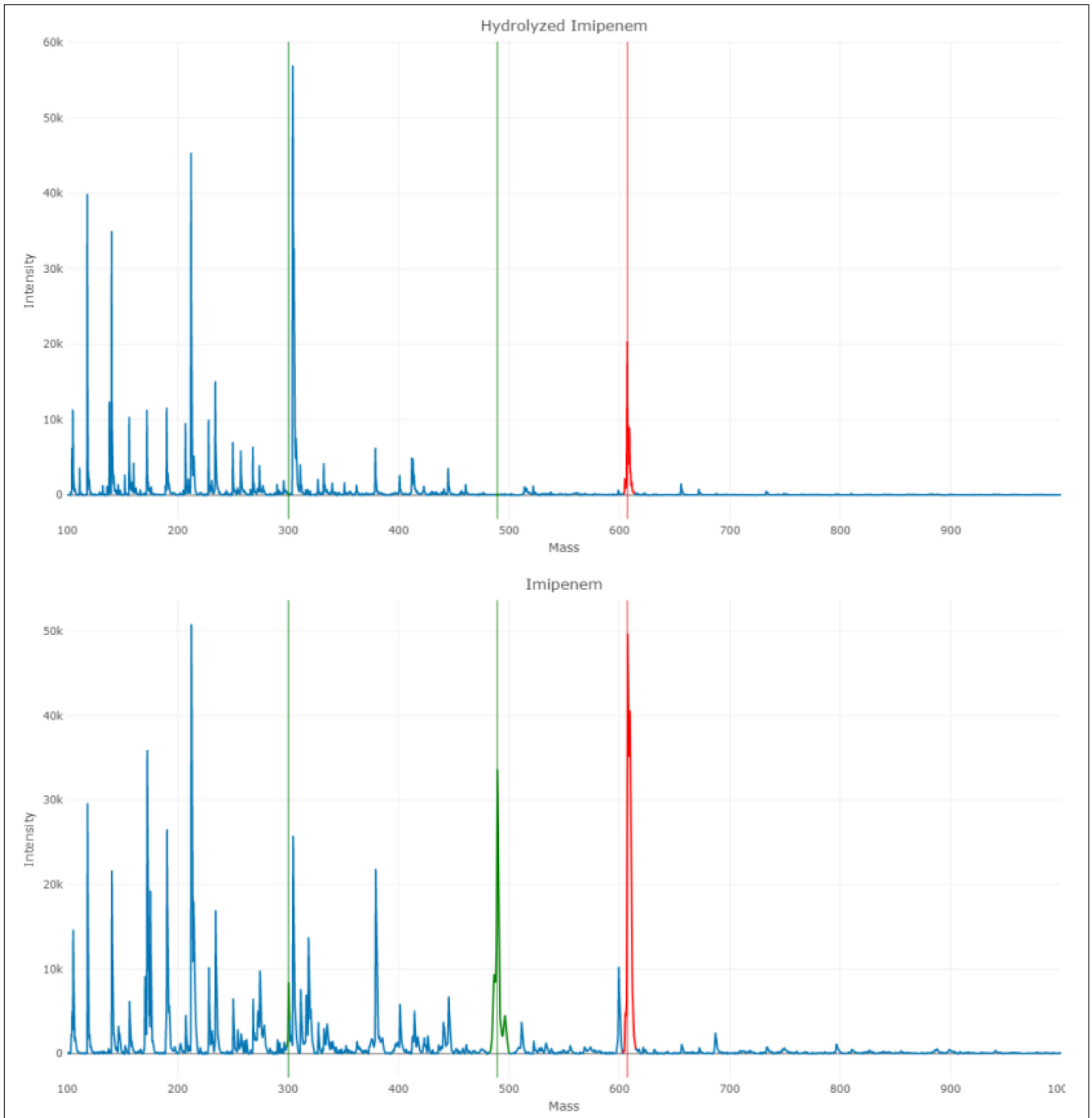


FIGURE 1 | Imipenem spectra. The hydrolyzed imipenem spectrum, in the first picture, represents no visible mass peaks for imipenem and only the internal standard, reserpine, is observed at 607 m/z. The later spectrum corresponds to imipenem, with the mass peaks at 300 and 489 m/z and the internal standard at 607 m/z.

(92.9–96.4); with minor errors representing 3.5% (2.4–5.3) and very major errors 1.5% (0.8–2.8) of the total. No major errors were found. The detailed results for each center are shown in **Table 2**. The best results were obtained in GM and in PZ, with 98.6% agreement, and the worst results were obtained in CA, with 88.9% agreement. Minor errors and very major errors ranged between, respectively, 1.4 and 8.3% and 1.4 and 5.6%. Regarding exclusively carbapenemase-producing isolates, the overall agreement rate was 95.5% (92.9–97.1, with minor errors representing 2.0% (1.0–3.9) and very major errors 2.5% (1.4–4.6) of the total.

The rates of agreement regarding the species are shown in **Table 3**. Values range from 91.6% (87.2–94.5) for *E. coli* to 100.0% (70.1–100.0) for *K. oxytoca*. The rates of agreement in

TABLE 2 | Agreement rates for each center in the intercenter evaluation.

Center[1]	Agreement rate[2] (*n* = 76)	Minor errors[3]	Very major errors[4]
GM	98.6%	1.4%	0%
RC	95.8%	1.4%	2.8%
RS	94.4%	4.2%	1.4%
VM	90.3%	8.3%	1.4%
SE	90.3%	8.3%	1.4%
CA	88.9%	8.3%	2.8%
AC	93.1%	6.9%	0%
PZ	98.6%	1.4%	0%
SL	91.7%	2.8%	5.6%
Total	94.9%	3.5%	1.5%

[1]The eight participating centers were Hospital General Universitario Gregorio Marañón (GM), Madrid; Hospital Universitario Ramón y Cajal (RC), Madrid; Hospital Universitario Reina Sofía (RS), Córdoba; Hospital Universitario Virgen Macarena (VM), Seville; Hospital Universitario Son Espases (SE), Palma; Hospital Universitario Central de Asturias (CA), Oviedo; University Hospital Plzeň (PZ), Plzeň and Cliniques Universitaires Saint-Luc (SL), Brussels, along with the reference center Complejo Hospitalario Universitario A Coruña (AC), A Coruña.
[2]Full agreement was considered when the MALDI-TOF assay provided an imipenem positive hydrolysis result for CPE or a negative hydrolysis result for non CPE.
[3]Minor errors were considered when the result of the MALDI-TOF assay was in the intermediate category and the isolate was either CPE or non CPE.
[4]Very major errors were considered when the result of the MALDI-TOF assay was negative for imipenem hydrolysis in CPE. No major errors were found in the study.

TABLE 3 | Agreement rates for different species in the intercenter evaluation for all centers participating in the study.

Species	Diagnosis	*N*[1]	%[2]	CI 95%[3]
K. pneumoniae	Agreement rate	329	96.2%	93.6–97.8%
	Minor errors	9	2.6%	1.4–4.9%
	Very major errors	4	1.2%	0.5–3.0%
E. coli	Agreement rate	206	91.6%	87.2–94.5%
	Minor errors	13	5.8%	3.4–9.6%
	Very major errors	6	2.7%	1.2–5.7%
E. cloacae	Agreement rate	71	98.6%	92.5–99.8%
	Minor errors	1	1.4%	0.2–7.5%
	Very major errors	0	0.0%	0.0–5.1%
K. oxytoca	Agreement rate	9	100.0%	70.1–100.0%
	Minor errors	0	0.0%	0.0–29.9%
	Very major errors	0	0.0%	0.0–29.9%

[1]Number of isolates included in each category.
[2]Percentage of isolates from each species that fit in each diagnosis category.
[3]The calculations are performed for a 95% confidence interval (CI).

relation to the resistance mechanisms are shown in **Table 4**. The rates of agreement ranged from 87.8% for VIM to 100% for K1. Regarding the accuracy of the MALDI-TOF method for detection of carbapenemases, the best classification rates were yielded for KPC-producing bacteria, with 98.9% of agreement relative to molecular methods.

Two isolates yielded most of the errors in the "intercenter evaluation," both are VIM-producing *E. coli*. The first one yielded minor and very major errors in 3 and in 1 out of the 9 centers, respectively. The second one yielded very major errors in 5 out of the 9 centers.

TABLE 4 | Agreement rates regarding the resistance mechanism obtained in the intercenter evaluation for all centers participating in the study.

Resistance mechanism	Diagnosis	*N*	%
Susceptible	Agreement rate	51	94.40%
	Minor errors	3	5.60%
	Very major errors	0	0.00%
CIT	Agreement rate	8	88.90%
	Minor errors	1	11.10%
	Very major errors	0	0.00%
SHV	Agreement rate	8	88.90%
	Minor errors	1	11.10%
	Very major errors	0	0.00%
CMY-2	Agreement rate	17	94.40%
	Minor errors	1	5.60%
	Very major errors	0	0.00%
FOX	Agreement rate	24	88.90%
	Minor errors	3	11.10%
	Very major errors	0	0.00%
K-1	Agreement rate	9	100.00%
	Minor errors	0	0.00%
	Very major errors	0	0.00%
CTX-M	Agreement rate	120	95.20%
	Minor errors	6	4.80%
	Very major errors	0	0.00%
KPC	Agreement rate	89	98.90%
	Minor errors	0	0.00%
	Very major errors	1	1.10%
IMP	Agreement rate	26	96.30%
	Minor errors	1	3.70%
	Very major errors	0	0.00%
NDM	Agreement rate	43	95.60%
	Minor errors	0	0.00%
	Very major errors	2	4.40%
VIM	Agreement rate	79	87.80%
	Minor errors	4	4.40%
	Very major errors	7	7.80%
OXA-48-like	Agreement rate	141	97.90%
	Minor errors	3	2.10%
	Very major errors	0	0.00%

Clinical Validation

The participating centers GM, RC, RS, PZ, and SL analyzed 20 isolates in the 2-month period; CA analyzed 19 isolates, SE analyzed 18 isolates and VM analyzed 10 isolates (**Supplementary Table 3**). The 147 isolates included were obtained from different clinical samples, urine (54), colonization samples (51), blood (16), wounds (8), respiratory samples (7), abscesses (5), catheter (4), and biopsies (2). Regarding the reference methods for detection of resistance, GM, CA, and SL used the O.K.N.V immunochromatographic test from CORIS BioConcept (Belgium); RC, PZ, RS, and SE used characterization by PCR and VM by WGS. The species most frequently detected were *K. pneumoniae* (67), followed at some distance by *E. cloacae* (Brackmann et al., 2020) and *E. coli* (Pitout et al., 2019). The carbapenemase enzyme most

frequently detected was OXA-48-like (69), followed by VIM (31), KPC (21), NDM (12), and IMP (5). The distribution of the carbapenemase enzymes among the different centers is shown in **Table 5**.

For statistical calculations, a total of 294 spectra were analyzed for duplicate measurements of 147 isolates. The overall agreement rate relative to the reference method for resistance detection was 92.5% (87.1–95.8), with 4.1% minor errors (1.9–8.6) and 3.4% very major errors (1.5–7.7) (**Table 6**). Thus, the sensitivity of the MALDI-TOF assay for detecting CPE was 93.9% (88.8–96.7). As no major errors were also found in the clinical validation stage the specificity of the methodology is 100% (70.1–100.0). The best results were obtained in VM and in PZ, with 100% of agreement and the worst results were obtained in SE, with 83.3% agreement. Minor and very major errors ranged between 0 and 10.0% and between 0 and 10%, respectively. Regarding the detection of the different carbapenemases enzymes, the agreements rates ranged from 88.4% for an OXA-48-like-producing isolate to 100% for KPC, NDM, and IMP. Minor and very major errors ranged between 0% (for VIM, KPC, NDM, and IMP) and 5.8% (for OXA-48-like).

TABLE 6 | greement rates for the different centers obtained in the clinical validation stage.

Center (*N*)[1]	Number of CPE[2]	Reference method[3]	Diagnosis		
			Agreement rate	Minor error	Very major error
GM (20)	19	IC	85.0%	5.0%	10.0%
RC (20)	20	PCR	95.0%	5.0%	0.0%
RS (20)	20	PCR	90.0%	10.0%	0.0%
VM (10)	10	WGS	100.0%	0.0%	0.0%
SE (18)	11	PCR	83.3%	5.6%	10.0%
CA (19)	19	IC	94.7%	5.3%	0.0%
PZ (20)	20	PCR	100.0%	0.0%	0.0%
SL (20)	19	IC	95.0%	0.0%	5.0%
Total (147)	138		92.5%	4.1%	3.4%

[1]Number of strains used for clinical validation in each center.
[2]Number of CPE confirmed among all the isolates suspected of having a carbapenemase enzyme by EUCAST screening criteria.
[3]The reference method used in each center for comparison with the results obtained with MALDI-TOF is described here. IC stands for the immunochromatographic test from CORIS BioConcept (Belgium), the rest of techniques were molecular based methods, like PCR and WGS.

DISCUSSION

MALDI-TOF MS is a simple, rapid procedure for detecting antimicrobial resistance that combines the universal advantages of phenotypic assays with the rapidity and accuracy of molecular assays. Different approaches have been developed for antimicrobial resistance detection by MALDI-TOF, including the quantitative growth assay (Idelevich et al., 2018) and the peak profiling methodology (Josten et al., 2014; Lau et al., 2014; Cordovana et al., 2018; Brackmann et al., 2020; Gato et al., 2021). However, none of these have been recommended by the EUCAST committee. To our knowledge, this is the first multicenter study evaluating the performance of MALDI-TOF MS for detecting carbapenemase activity in Enterobacterales in clinical microbiology laboratories. At the "software design" stage, we realized that our in-house procedure with data analysis in the Clover MS data analysis software (Clover BioSoft, Spain) yielded stronger agreement relative to the results obtained with the MBT STAR®-Carba IVD Kit (Bruker Daltonik) in the STAR®-BL IVD software (96.30% versus 92.00%). Our procedure also provides several other advantages. First, it is commercially independent from the manufacturer of the MALDI-TOF instrument, so that different spectra formats are compatible. In this study, all collaborators used Bruker Daltonik equipment; however, the data analysis can also be performed from mass spectra originated with bioMérieux mass spectrometer devices (Marcy-l'Étoile, France) an others, without the need for further modifications (data not shown). The software automatically converts any mass spectra format in a common format used for interpretation of spectra without any previous knowledge from the user. This is a particularly useful feature, as multiple instrument platforms have emerged since the universal integration of MALDI-TOF in microbiology laboratories. Furthermore, the software is available online and the information generated can thus be accessed anywhere at any time. This is a huge advantage for clinical use, as MALDI-TOF instruments are used for high-throughput screening in microbiology laboratories for identification purposes, so the online application allows access from any computer in the lab, releasing the MALDI-TOF computer for identification and delaying the resistance analysis to any suitable moment. The online application also enables networking with colleagues in different centers, so that information can be shared and advice can even be requested

TABLE 5 | Distribution of resistance mechanism in relation to the centers in the clinical validation.

Center[1] (*N*[2])	Resistance mechanism					
	OXA	VIM	KPC	NDM	IMP	Non-carbapenem producers
GM (20)	20	0	0	0	0	0
RC (20)	9	7	4	0	0	0
RS (20)	6	4	6	0	4	0
VM (10)	5	2	2	0	1	0
SE (18)	1	10	0	0	0	7
CA (19)	16	2	0	0	0	1
PZ (20)	5	4	6	5	0	0
SL (20)	7	2	3	7	0	1
Total (147)	69	31	21	12	5	9

[1]Hospital General Universitario Gregorio Marañón (GM), Madrid; Hospital Universitario Ramón y Cajal (RC), Madrid; Hospital Universitario Reina Sofía (RS), Córdoba; Hospital Universitario Virgen Macarena (VM), Seville; Hospital Son Espases (SE), Palma; Hospital Universitario Central de Asturias (CA), Oviedo; Plzeň University Hospital (PZ), Plzeň and Cliniques Universitaires Saint-Luc (SL), Brussels.
[2]Number of strains used for clinical validation in each center.

from reference centers, as it is currently very common for hospitals to unify forces and work together as bigger working areas that share patients and information. Besides, the software includes the database of spectra introduced and saved by the user, that can be shared upon permission with any other user of the software. In the "intercenter evaluation" results obtained by the different centers were compared with the molecular characterization of the isolates. In the "clinical validation" results were not compared to a unique standard, but to different and independent validated techniques used in real clinical situations for carbapenemase detection in Enterobacterales, facilitating the user familiarization with the developed methodology In the "intercenter evaluation," we have demonstrated that the procedure developed is highly reproducible, as agreements rates above 90% were reached in all centers, except one, with a range of values across centers of less than 10%. The inexperience of the user plus a non-fully adjusted equipment by an experienced technician could be the reasons of the lower rates. Moreover, in the "clinical validation" stage, the levels of agreement relative to the reference techniques applied in each laboratory were very similar to those obtained in the "intercenter evaluation," i.e., 92.5% versus 94.9%, thus validating the technique for use in clinical practice. Minor errors did not exceed 5% and very major errors were less than 3.5% in the real clinical setting, thus preventing further unnecessary testing of isolates.

In the "intercenter evaluation," study of the rates of agreement regarding the species did not indicate any significant differences, with a very similar level of accuracy for different bacteria. Regarding the "clinical validation" stage, no statistical analysis was possible, because of the small number of isolates of each bacterial species.

Regarding the agreement rates in the "clinical validation" stage for the different carbapenemase enzymes, detection of OXA-48-like-producing isolates was slightly poorer than for the other CPE isolates, i.e., 88.4% versus 100%. However, in the "intercenter evaluation," detection of VIM-producing isolates was poorest, detected at a rate of 87.8%, with OXA-48-like-producing isolates detected with an accuracy rate of 97.9%. Possible reasons for these differences include the presence of different copies of the resistance plasmid and the isolates belonging to different, epidemiologically unrelated sequence types expressing different degrees of susceptibility to carbapenems. However, more extensive evaluation should be performed in order to establish the reasons for these differences.

Neither in the "intercenter evaluation" nor in the "clinical validation" stage, major errors were produced, thus preventing false-positive detection of AmpC or ESBL-producers, errors commonly found in phenotypic methods of carbapenemase detection (Hrabák et al., 2014; European Center for Disease Prevention and Control, 2020).

Regarding the results bracket among the different centers, 88.9–198.6% in the "intercenter evaluation and 83.3–100% in the "clinical validation" stages, the variance is higher in the second part. This can be attributable to the biological differences among isolates and also in the number of CPE confirmed. Not all suspected isolates of having a carbapenemase by EUCAST screening criteria, were confirmed by the reference method used by the respective laboratory. For example, the center having the lowest accuracy data, SE, had only 11 CPE, so errors have higher impact in the final accuracy percentage, than for example in RC, PZ, or SL with 20 CPE analyzed. As only one center used as reference method WGS, false positive results obtained by MALDI-TOF could be wrongly attributable, really corresponding to true carbapenemase isolates rarely detected as IMI, SME, and FRI… However, as all participating laboratories have WGS tools (although not used routinely), it is not likely to have happened, as repeated isolates with reduced susceptibility to carbapenems with no carbapenemase found by their routine diagnostics methodologies would have been submitted to full sequencing, so no outbreak should have been missed during the study. The common slight variations can be allocated to differences in the level of expertise of the users, previous or not experience using MALDI-TOF, or even slight differences in the adjustments of equipments. However, all users accomplished the task without on site training, just with the information provided in this paper, so we proved that this developed methodology is ready to be performed in any microbiology laboratory with minimal equipment and no previous knowledge. As a limitation in our study, we must recognize the low number of clinical isolates in the "clinical validation," thus concluding that the method is promising, but requires a more robust and comprehensive clinical validation.

The MALDI-TOF hydrolysis assay has the advantage of being able to detect carbapenemase activity regardless of the enzyme produced, including novel enzymes emerging at any given time which are not detected by predefined PCR tools usually used in diagnostics (Yoon et al., 2020). In addition, false positive results can also occur when PCR tools are used, due to the appearance of novel variants which lack carbapenemase activity, such as KPC-28 and OXA-163 (Dortet and Naas, 2017; Oueslati et al., 2019), or with reduced carbapenemase activity, such as OXA-232 and OXA-244 (Potron et al., 2013; Hoyos-Mallecot et al., 2017). MALDI-TOF MS measures carbapenemase activity, acting like a rapid phenotypic method. However, additional techniques are required for precise identification of the carbapenemase class. Recent publications have overcome this issue, demonstrating the potential use of MALDI-TOF MS for carbapenemase classification using different combinations of β-lactamase inhibitors (Carvalhaes et al., 2018; Oviaño et al., 2020). The requirement for testing the isolates from blood agar cultures is a limitation of our study, as no other culture media has been tested. However, in our experience the hydrolysis assay by MALDI-TOF works very similarly using bacterial colonies from other type of culture media as McConkey agar or Mueller-Hinton agar (data not shown). Also, the need of fresh controls could be observed as a limitation of the study, however, most laboratories that have an antimicrobial resistance detection section will have this controls ready to use for different findings.

In comparison with biochemical assays, MALDI-TOF MS has the advantage of overcoming the subjective visual interpretation,

particularly in case of weak positive reactions as in OXA-48-like enzymes, where the color change is not as evident. In comparison with other in-house methods, the major advantage of the method is the standardized processing, with no need to establish a particular number of cells of the isolate for analysis and no subsequent extraction (Hrabák et al., 2011; Monteferrante et al., 2016). The excellent performance of the buffer enables an all-in-one (Oviaño et al., 2016), simplified procedure, with sensitivity and specificity values in the same range.

The use of our in-house procedure is more labor intensive than with the MBT STAR®-Carba IVD Kit (Bruker Daltonik), as it requires preparation of specific reagents, like the antibiotic solution, matrix and calibration standard, making in principle, the introduction of the method in routine less straightforward. This is why the working protocol with the standardized procedure and educational video for software management is so useful and helps in the familiarization of the methodology (**Supplementary Data Sheet 1** and **Supplementary Video 1**). The different staff performing the assay were not dedicated or expert personnel, but was the same person in each center during the entire study. The preparation of reagents consists in simple steps of weighing and dilution, techniques commonly used in microbiology laboratories that take about 10–20 min to accomplish, with an extra 30 min of incubation. Besides, in the future automated machines can be programmed for preparation of MALDI-TOF reagents in the users way, as there are different protocols for DNA/RNA extraction and for PCR in molecular techniques. Comparing the turnaround time for both techniques, it is quite similar, with a short delay of for the preparation of reagents in our developed in-house procedure, with the additional advantages in the methodology developed herein of a very low cost (∼ 1 euro) and the possibility of adjusting the procedure for any antibiotic (Oviaño et al., 2017).

CONCLUSION

MALDI-TOF MS is an outstanding tool for the rapid detection of carbapenemase activity in Enterobacterales in clinical microbiology laboratories. The use of a simple in-house procedure with online software allows results to be obtained within 30–60 min, making MALDI-TOF a rapid phenotypic technique suitable for routine screening of carbapenemases in diagnostics and thus facilitating the early implementation of appropriate antimicrobial therapy.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

MO: contributions to the conception or design of the work. EG, AA, MArr, MArt, JC, AC, KC, IPC, CE, JF, JJ, PL, GM, XM, and PP-P: contributions to the acquisition and analysis. BR-S, RC, JH, LM, LM-M, AO, ÁP, AV, and GB: contributions to the drafting the work or revising it critically for important intellectual content. GB and MO: provide approval for publication of the content. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Fondo de Investigación Sanitaria (grant numbers PI20/00686 to MO and PI18/00860 to GB) integrated in the Plan Nacional de I+D and funded by the Instituto de Salud Carlos III (ISCIII). The results of this work have been funded by the Spanish Network of Research in Infectious Diseases (REIPI), N° RD16/0016/0006 and RD16/0016/0011, integrated in the National Plan for Scientific Research, Development and Technological Innovation 2013–2016, and funded by the ISCIII-General Subdirection of Assesment and Promotion of the Research – European Regional Development Fund (FEDER) "A way of making Europe."

ACKNOWLEDGMENTS

We thank the "Unidade de Apoio á Investigación" for the statistical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.789731/full#supplementary-material

REFERENCES

Anantharajah, A., Tossens, B., Olive, N., Kabamba-Mukadi, B., Rodriguez-Villalobos, H., and Verroken, A. (2019). Performance evaluation of the MBT STAR§ -Carba IVD assay for the detection of carbapenemases With MALDI-TOF MS. *Front. Microbiol.* 10:1413. doi: 10.3389/fmicb.2019.01413

Brackmann, M., Leib, S. L., Tonolla, M., Schürch, N., and Wittwer, M. (2020). Antimicrobial resistance classification using MALDI-TOF-MS is not that easy: lessons from vancomycin-resistant *Enterococcus faecium*. *Clin. Microbiol. Infect.* 26, 391–393. doi: 10.1016/j.cmi.2019.10.027

Burckhardt, I., and Zimmermann, S. (2011). Using matrix-assisted laser desorption ionization–time of flight mass spectrometry to detect carbapenem resistance within 1 to 2.5 hours. *J. Clin. Microbiol.* 49, 3321–3324. doi: 10.1128/JCM.00287-11

Carvalhaes, C. G., Ramos, A. C., Oliveira, L. C. G., Juliano, M. A., and Gales, A. C. (2018). Rapid detection of ceftazidime/avibactam resistance by MALDI-TOF MS. *J. Antimicrob. Chemother.* 73, 2579–2582. doi: 10.1093/jac/dky196

Clark, A. E., Kaleta, E. J., Arora, A., and Wolk, D. M. (2013). Matrix-assisted laser desorption ionization-time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. *Clin. Microbiol. Rev.* 26, 547–603. doi: 10.1128/CMR.00072-12

Cordovana, M., Abdalla, M., and Ambretti, S. (2020). Evaluation of the MBT STAR-Carba assay for the detection of carbapenemase production in *Enterobacteriaceae* and Hafniaceae with a large collection of routine isolates

from plate cultures and patient-derived positive blood cultures. *Microb. Drug Resist.* 26, 1298–1306.

Cordovana, M., Kostrzewa, M., Sóki, J., Witt, E., Ambretti, S., and Pranada, A. B. (2018). *Bacteroides fragilis*: a whole MALDI-based workflow from identification to confirmation of carbapenemase production for routine laboratories. *Anaerobe* 54, 246–253.

Decousser, J. W., Poirel, L., Desroches, M., Jayol, A., Denamur, E., and Nordmann, P. (2015). Failure to detect carbapenem-resistant *Escherichia coli* producing OXA-48-like using the Xpert Carba-R assay§. *Clin. Microbiol. Infect.* 21, e9–e10. doi: 10.1016/j.cmi.2014.09.006

Dortet, L., and Naas, T. (2017). Noncarbapenemase OXA-48 variants (OXA-163 and OXA-405) falsely detected as carbapenemases by the β-Carba Test. *J. Clin. Microbiol.* 55, 654–655. doi: 10.1128/JCM.02086-16

Dortet, L., Jousset, A., Sainte-Rose, V., Cuzon, G., and Naas, T. (2016). Prospective evaluation of the OXA-48 K-SeT assay, an immunochromatographic test for the rapid detection of OXA-48-type carbapenemases. *J. Antimicrob. Chemother.* 71, 1834–1840.

European Center for Disease Prevention and Control (2020). *The European Surveillance System Antimicrobial Resistance (AMR) Reporting Protocol 2020*. Solna Municipality: European Center for Disease Prevention and Control.

Gato, E., Constanso, I. P., Rodiño-Janeiro, B. K., Guijarro-Sánchez, P., Alioto, T., Arroyo, M. J., et al. (2021). Occurrence of the p019 gene in the blaKPC-harboring plasmids: adverse clinical impact for direct tracking of KPC-producing *Klebsiella pneumoniae* by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J. Clin. Microbiol.* 59:e0023821.

Hoyos-Mallecot, Y., Naas, T., Bonnin, R. A., Patino, R., Glaser, P., Fortineau, N., et al. (2017). OXA-244-producing *Escherichia coli* isolates, a challenge for clinical microbiology laboratories. *Antimicrob. Agents Chemother.* 61:e00818-17. doi: 10.1128/AAC.00818-17

Hrabák, J., Chudáèková, E., and Papagiannitsis, C. C. (2014). Detection of carbapenemases in *Enterobacteriaceae*: a challenge for diagnostic microbiological laboratories. *Clin. Microbiol. Infect.* 20, 839–853.

Hrabák, J., Walková, R., Studentová, V., Chudácková, E., and Bergerová, T. (2011). Carbapenemase activity detection by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J. Clin. Microbiol.* 49, 3222–3227.

Idelevich, E. A., Sparbier, K., Kostrzewa, M., and Becker, K. (2018). Rapid detection of antibiotic resistance by MALDI-TOF mass spectrometry using a novel direct-on-target microdroplet growth assay. *Clin. Microbiol. Infect.* 24, 738–743. doi: 10.1016/j.cmi.2017.10.016

Josten, M., Dischinger, J., Szekat, C., Reif, M., Al-Sabti, N., Sahl, H. G., et al. (2014). Identification of agr-positive methicillin-resistant *Staphylococcus aureus* harbouring the class A mec complex by MALDI-TOF mass spectrometry. *Int. J. Med. Microbiol.* 304, 1018–1023. doi: 10.1016/j.ijmm.2014.07.005

Lasserre, C., De Saint Martin, L., Cuzon, G., Bogaerts, P., Lamar, E., Glupczynski, Y., et al. (2015). Efficient detection of carbapenemase activity in *Enterobacteriaceae* by matrix-assisted laser desorption ionization-time of flight mass spectrometry in less than 30 minutes. *J. Clin. Microbiol.* 53, 2163–2171. doi: 10.1128/JCM.03467-14

Lau, A. F., Wang, H., Weingarten, R. A., Drake, S. K., Suffredini, A. F., Garfield, M. K., et al. (2014). 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.* 52, 2804–2812.

Monteferrante, C. G., Sultan, S., Ten Kate, M. T., Dekker, L. J., Sparbier, K., Peer, M., et al. (2016). Evaluation of different pretreatment protocols to detect accurately clinical carbapenemase-producing *Enterobacteriaceae* by MALDI-TOF. *J. Antimicrob. Chemother.* 71, 2856–2867. doi: 10.1093/jac/dkw208

Oteo, J., Pérez-Vázquez, M., Bautista, V., Ortega, A., Zamarrón, P., Saez, D., et al. (2016). The spread of KPC-producing *Enterobacteriaceae* in Spain: WGS analysisof the emerging high-risk clones of *Klebsiella pneumoniae* ST11/KPC-2, ST101/KPC-2 and ST512/KPC-3. *J. Antimicrob. Chemother.* 71, 3392–3399.

Oteo, J., Saez, D., Bautista, V., Fernández-Romero, S., Hernández-Molina, J. M., Pérez-Vázquez, M., et al. (2013). Spanish collaborating group for the antibiotic resistance surveillance program. Carbapenemase-producing *enterobacteriaceae* in Spain in 2012. *Antimicrob. Agents Chemother.* 57, 6344–6347.

Oueslati, S., Iorga, B. I., Tlili, L., Exilie, C., Zavala, A., Dortet, L., et al. (2019). Unravelling ceftazidime/avibactam resistance of KPC-28, a KPC-2 variant lacking carbapenemase activity. *J. Antimicrob. Chemother.* 74, 2239–2246. doi: 10.1093/jac/dkz209

Oviaño, M., and Bou, G. (2018). Matrix-assisted laser desorption ionization-time of flight mass spectrometry for the rapid detection of antimicrobial resistance mechanisms and beyond. *Clin. Microbiol. Rev.* 32:e00037-18. doi: 10.1128/CMR.00037-18

Oviaño, M., Gato, E., and Bou, G. (2020). Rapid detection of KPC-producing enterobacterales susceptible to imipenem/relebactam by using the MALDI-TOF MS MBT STAR-Carba IVD assay. *Front. Microbiol.* 11:328. doi: 10.3389/fmicb.2020.00328

Oviaño, M., Gómara, M., Barba, M. J., Revillo, M. J., Barbeyto, L. P., and Bou, G. (2017). Towards the early detection of β-lactamase-producing *Enterobacteriaceae* by MALDI-TOF MS analysis. *J. Antimicrob. Chemother.* 72, 2259–2262. doi: 10.1093/jac/dkx127

Oviaño, M., Rodicio, M. R., Heinisch, J. J., Rodicio, R., Bou, G., and Fernández, J. (2019). Analysis of the degradation of broad-spectrum cephalosporins by OXA-48-producing *Enterobacteriaceae* using MALDI-TOF MS. *Microorganisms* 7:614. doi: 10.3390/microorganisms7120614

Oviaño, M., Sparbier, K., Barba, M. J., Kostrzewa, M., and Bou, G. (2016). Universal protocol for the rapid automated detection of carbapenem-resistant Gram-negative bacilli directly from blood cultures by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF/MS). *Int. J. Antimicrob. Agents* 48, 655–660. doi: 10.1016/j.ijantimicag.2016.08.024

Papagiannitsis, C. C., Študentová, V., Izdebski, R., Oikonomou, O., Pfeifer, Y., Petinaki, E., et al. (2015). Matrix-assisted laser desorption ionization-time of flight mass spectrometry meropenem hydrolysis assay with NH_4HCO_3, a reliable tool for direct detection of carbapenemase activity. *J. Clin. Microbiol.* 53, 1731–1735. doi: 10.1128/JCM.03094-14

Pitout, J. D. D., Peirano, G., Kock, M. M., Strydom, K. A., and Matsumura, Y. (2019). The global ascendency of OXA-48-Type carbapenemases. *Clin. Microbiol. Rev.* 33:e00102-19. doi: 10.1128/CMR.00102-19

Potron, A., Rondinaud, E., Poirel, L., Belmonte, O., Boyer, S., Camiade, S., et al. (2013). Genetic and biochemical characterisation of OXA-232, a carbapenem-hydrolysing class D β-lactamase from *Enterobacteriaceae*. *Int. J. Antimicrob. Agents* 41, 325–329. doi: 10.1016/j.ijantimicag.2012.11.007

Rodríguez-Baño, J., Gutiérrez-Gutiérrez, B., Machuca, I., and Pascual, A. (2018). Treatment of infections caused by extended-spectrum-β- Lactamase-, AmpC-, and carbapenemase-producing *Enterobacteriaceae*. *Clin. Microbiol. Rev.* 31:e00079-17. doi: 10.1128/CMR.00079-17

U.S. Centers for Disease Control and Prevention (2019). *Antibiotic Resistance Threats in the United States, 2019*. Atlanta, GA: U.S. Centers for Disease Control and Prevention.

Yoon, E. J., Choi, Y. J., Park, S. H., Shin, J. H., Park, S. G., Choi, J. R., et al. (2020). A novel KPC variant KPC-55 in *Klebsiella pneumoniae* ST307 of reinforced meropenem-hydrolyzing activity. *Front. Microbiol.* 11:561317. doi: 10.3389/fmicb.2020.561317

Conflict of Interest: MArr, JJ, GM, and LM were employed by company Clover BioSoft.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 18 February 2022
doi: 10.3389/fmicb.2021.804484

Combination of MALDI-TOF Mass Spectrometry and Machine Learning for Rapid Antimicrobial Resistance Screening: The Case of *Campylobacter* spp.

Maureen Feucherolles[1], Morgane Nennig[2], Sören L. Becker[3,4,5], Delphine Martiny[6,7], Serge Losch[8], Christian Penny[1,9], Henry-Michel Cauchie[1]* and Catherine Ragimbeau[2]*

[1] Environmental Research and Innovation (ERIN) Department, Luxembourg Institute of Science and Technology, Belval, Luxembourg, [2] Laboratoire National de Santé, Epidemiology and Microbial Genomics, Dudelange, Luxembourg, [3] Institute of Medical Microbiology and Hygiene, Saarland University, Homburg, Germany, [4] Swiss Tropical and Public Health Institute, Basel, Switzerland, [5] University of Basel, Basel, Switzerland, [6] National Reference Centre for Campylobacter, Laboratoire des Hôpitaux Universitaires de Bruxelles-Universitaire Laboratorium Brussel (LHUB-ULB), Brussels, Belgium, [7] Université de Mons (UMONS), Mons, Belgium, [8] Laboratoire de Médecine Vétérinaire de l'Etat, Dudelange, Luxembourg, [9] Chambre des Députés du Grand-Duché de Luxembourg, Parliamentary Research Service, Luxembourg, Luxembourg

OPEN ACCESS

Edited by:
Antonella Lupetti, University of Pisa, Italy

Reviewed by:
Miriam Cordovana, Bruker Daltonik GmbH, Germany
Marina Oviaño, A Coruña University Hospital Complex (CHUAC), Spain

***Correspondence:**
Maureen Feucherolles
feucherolles.maureen@gmail.com
Henry-Michel Cauchie
henry-michel.cauchie@list.lu

Specialty section:
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

Received: *29 October 2021*
Accepted: *28 December 2021*
Published: *18 February 2022*

Citation:
Feucherolles M, Nennig M, Becker SL, Martiny D, Losch S, Penny C, Cauchie H-M and Ragimbeau C (2022) Combination of MALDI-TOF Mass Spectrometry and Machine Learning for Rapid Antimicrobial Resistance Screening: The Case of Campylobacter spp. Front. Microbiol. 12:804484. doi: 10.3389/fmicb.2021.804484

While MALDI-TOF mass spectrometry (MS) is widely considered as the reference method for the rapid and inexpensive identification of microorganisms in routine laboratories, less attention has been addressed to its ability for detection of antimicrobial resistance (AMR). Recently, some studies assessed its potential application together with machine learning for the detection of AMR in clinical pathogens. The scope of this study was to investigate MALDI-TOF MS protein mass spectra combined with a prediction approach as an AMR screening tool for relevant foodborne pathogens, such as *Campylobacter coli* and *Campylobacter jejuni*. A One-Health panel of 224 *C. jejuni* and 116 *C. coli* strains was phenotypically tested for seven antimicrobial resistances, i.e., ciprofloxacin, erythromycin, tetracycline, gentamycin, kanamycin, streptomycin, and ampicillin, independently, and were submitted, after an on- and off-plate protein extraction, to MALDI Biotyper analysis, which yielded one average spectra per isolate and type of extraction. Overall, high performance was observed for classifiers detecting susceptible as well as ciprofloxacin- and tetracycline-resistant isolates. A maximum sensitivity and a precision of 92.3 and 81.2%, respectively, were reached. No significant prediction performance differences were observed between on- and off-plate types of protein extractions. Finally, three putative AMR biomarkers for fluoroquinolones, tetracyclines, and aminoglycosides were identified during the current study. Combination of MALDI-TOF MS and machine learning could be an efficient and inexpensive tool to swiftly screen certain AMR in foodborne pathogens, which may enable a rapid initiation of a precise, targeted antibiotic treatment.

Keywords: MALDI-TOF MS, antimicrobial resistance screening, AMR, machine learning, *Campylobacter*, diagnostics

INTRODUCTION

Antimicrobial susceptibility testing (AST) is a key technology in diagnostic microbiology and is essential for a targeted treatment and to limit the widespread use of broad-spectrum antibiotics. Over the past decades, many improvements have helped to accelerate, standardize, and harmonize testing facilities, e.g., through the implementation of automated and semi-automated devices combining identification and AST (e.g., Vitek 2®), using optical systems for measuring changes in bacterial growth and determining antimicrobial susceptibility, and using rapid diagnostic tests for same-day AST results (Mitchell and Alby, 2017; Benkova et al., 2020; Roth et al., 2021). In a concern for harmonization, disk-diffusion and microdilution antibiograms, recommended by the European committee on antimicrobial susceptibility testing (EUCAST, human medicine) or the European food safety authority (EFSA, veterinary medicine), are still the reference methods for determination of antimicrobial resistances (AMR). These tests are based on bacterial growth, requiring between 16 and 24 h for rapid growing pathogens and longer for fastidious pathogens (e.g., mycobacteria and *Helicobacter pylori*) (Barlam et al., 2016; Arena et al., 2017). Results are usually qualitative and classed into categories, i.e., susceptible or resistant, depending on the breakpoint calibrated by the EUCAST, or expressed as minimum inhibitory concentration (MIC) (Benkova et al., 2020). While these conventional methods are effective, they are cumbersome, time-consuming, and do not enable the rapid choice of an effective targeted anti-infective treatment. Yet, development of "fast microbiology" technologies or rapid diagnostic tests, including Matrix assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS), results in the improvement of the antimicrobial stewardship by decreasing the "patient–physician" workflow before treatment (Bookstaver et al., 2017; Mangioni et al., 2019).

MALDI-TOF MS is a soft-ionized mass spectrometry method developed as an analytical tool to identify and understand the structure of unknown biomolecules (Gibson and Costello, 2000). In an evolving field, this automatic technique became the reference method for identifying microorganisms such as bacteria (Clark et al., 2013; Singhal et al., 2015), mycobacteria (Rodriguez-Granger et al., 2018; Rotcheewaphan et al., 2019) and fungi (Florio et al., 2018; Robert et al., 2021). The resolution power of the system operates at the species level and even at sub-species level for a number of pathogens in clinical microbiology (Fall et al., 2015; Feucherolles et al., 2021). It is a fast and cost-efficient process, with a positive impact on public health analytical pipelines (Ge et al., 2017; Rodríguez-Sánchez et al., 2019). Identification of other organisms, like protozoa (Del Chierico et al., 2016), helminths (Bredtmann et al., 2017; Feucherolles et al., 2019b; Sy et al., 2021; Wendel et al., 2021), viruses (Iles et al., 2020; Rybicka et al., 2021), and arthropods (Tahir et al., 2017; Boucheikhchoukh et al., 2018; Tandina et al., 2018), is also feasible in a research context. However, only the routine identification part of the diagnostics workflow is currently carried out by MALDI-TOF MS.

Over the last 5 years, machine learning (ML), a subset of artificial intelligence, has gained interest in many areas of research pertaining to an improved diagnosis of diseases (e.g., cancer detection, infectious diseases, etc.) (Caballé et al., 2020; Goodswen et al., 2021; Nami et al., 2021). This popularity is greatly explained by the current era, where large daily amounts of data are being collected digitally, known as big data, which are requiring new approaches to investigate it. Mass spectra are routinely generated by MALDI-TOF MS and most of the time not exploited for additional analysis beyond the sole identification of microorganisms. Even if several reports highlighted successful applications of MALDI-TOF MS for detection of bacterial AMR, by the presence of specific biomarkers (Feucherolles et al., 2019a; Oviaño and Bou, 2019; Yoon and Jeong, 2021) identified by classical statistical methods, there is still a mine of information encrypted in the mass spectra. More recently, a growing number of reports combining MALDI-TOF mass spectrometry and ML have shown promising results for clinical big data problems, such as AMR screening (Weis et al., 2020a,b). The majority of these studies used pathogens such as *Staphylococcus aureus* and the β-lactam antibiotic family (Sogawa et al., 2017; Wang et al., 2018; Tang et al., 2019). Therefore, there are very few published data concerning other relevant clinical or foodborne pathogens or antimicrobials such as the quinolones (e.g., ciprofloxacin) and macrolides (e.g., erythromycin and azithromycin) (Sabença et al., 2020; Sousa et al., 2020). However, macrolides and quinolones are frontline antibiotics used to treat severe infectious gastroenteritis and categorized by the World Health Organization (WHO) as critically important in human medicine (WHO, 2019).

Campylobacteriosis, mainly caused by *C. jejuni* and *C. coli*, is the main global cause of bacterial gastroenteritis in humans (Chlebicz and Śliżewska, 2018). Likewise, 10.9 and 0.6% of *C. coli* and *C. jejuni,* respectively, isolated from humans were multi-resistant to ciprofloxacin, erythromycin, tetracycline, and gentamycin in 2019 (EFSA and ECDC, 2021). In food-producing animals, 26.9% of *C. coli* isolated from calves were resistant to at least three of the previously cited antimicrobials. MALDI-TOF MS already has been applied for proteo-typing of *C. coli*, *C. fetus*, and more recently for *C. concisus* genomospecies (Emele et al., 2019a,b; On et al., 2021). Also, its ability to distinguish β-lactam-resistant strains from sensitive ones by pre-processing mass spectra before analysis was reported (Penny et al., 2016). However, there are no published reports concerning the direct application of the mass spectrometry and ML for direct prediction of AMR in *Campylobacter* spp.

Therefore, the aim of this study is to show that MALDI-TOF MS combined with an ML approach could be a useful tool for a fast and precise AMR screening of relevant foodborne pathogens, such as *C. coli* and *C. jejuni*. While campylobacteriosis is mainly self-limiting and do not require specific antibiotherapy, such a combination strategy may aid to swiftly prescribe a definitive antimicrobial therapy and therefore limit an empirical broad-spectrum strategy for other pathogens. ML prediction based on protein mass spectra will be investigated at the species-specific and antibiotic resistance level. The impact of different protein extraction methods, i.e., on- and off-plate extraction, on resistance predictions will also be considered.

MATERIALS AND METHODS

Campylobacter Collection

Strains

A One-Health collection of 224 *C. jejuni* and 116 *C. coli* isolates, obtained from humans (n = 226), in environmental samples, i.e., surface water (n = 33), and animals including wild life: raccoons (n = 8), wild birds (n = 17), and cattle, i.e., bovine (n = 20), pig (n = 1), and poultry (n = 35), were used in the current study.

Antimicrobial resistances patterns were established by disk diffusion antibiograms for fluroquinolones [ciprofloxacin (Cip, 5 μg)], macrolides [erythromycin (Ery, 15 μg)], tetracyclines [tetracycline (Tet, 30 μg)], aminoglycosides [gentamycin (Gent, 10 μg), kanamycin (Kana, 30 μg), Streptomycin (Strep, 10 μg)], and β-lactams [ampicillin (Amp, 10 μg)] following the French Microbiology Society (SFM) and EUCAST recommendations (Recommendations 2020 v1.1 April) resulting in patterns addressed in **Table 1**. For antibiotics not described for *Campylobacter* spp., i.e., kanamycin and streptomycin, EUCAST recommendation for the *Enterobacterales* group was applied. The latter was added to the study based on ResFinder analysis by using Whole Genome Sequencing (WGS) data (Bortolaia et al., 2020). The Lys43Arg mutation in the *rspL* gene as well as *ant*(6) and *aadE* genes and conferring the streptomycin resistance were detected (Olkkola et al., 2010; Fabre et al., 2018). Likewise, the *aph*(3) gene conferring among other kanamycin resistance was detected (Fabre et al., 2018). The phenotypic details of the collection are described in **Supplementary File 1**.

Growth Conditions

All strains were inoculated on chocolate agar plates (Thermo Scientific, Waltham, MA, United States) with -80°C stock suspension stored in FBP medium complemented with *Campylobacter* growth supplement (Thermo Fisher Scientific), and incubated for 48 ± 2 h at 42°C under micro aerobic conditions using CampyGen 2.5 L gas packs (Thermo Fisher Scientific).

Matrix Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry Analysis

Sample Preparation

For every biological assay, an off- and on-plate extraction and a direct deposit were performed. For the off-plate or also known as ethanol/formic acid protein extraction (EtOH/ACN), bacteria were suspended in 300 μl milliQ water and 900 μl absolute ethanol (Merck, Darmstadt, Germany). The mix was centrifuged for a further 2 min and the residual ethanol was discarded. A total of 25 μl for both 70% (v/v) formic acid (Merck, Darmstadt, Germany) and acetonitrile (Merck) were mixed up to the dry pellet. A final centrifugation was performed, and then 1 μl of supernatant was spotted onto a one-use MALDI Biotarget (96 targets; Bruker Daltonics GmbH, Bremen, Germany). For the formic acid on-plate extraction (FA), a smear of a bacteria colony is directly carried out on the biotarget and then overlayed with a 1 μl 70% formic acid. For the direct deposit, a bacteria colony is directly streaked on the biotarget. For all deposits and extractions, as soon as the sample was dried, the spot was overlaid with 1 μl of portioned HCCA matrix solution (Bruker Daltonics GmbH) prepared with standardized acetonitrile 50%, water 47.5%, and trifluoroacetic acid 2.5% solution (Sigma-Aldrich, Saint Louis, MO, United States). Bruker bacterial test standard (BTS) was used for an external calibration of the apparatus.

For each method of extraction, three independent cultures (biological replicates) on three different days (reproducibility) were performed. Each biological replicate was spotted thrice (technical replicates) on the same day (repeatability), resulting in nine spectra per isolate.

Data Acquisition

MALDI-TOF MS analysis was performed using a Biotyper Microflex LT/SH (Bruker Daltonics GmbH) by using the AutoXecute acquisition method (MBT_AutoX) in FlexControl software v3.4., with a 2–20 kDa mass-to-charge ratio (m/z) range in a positive linear mode. Before measurement, the system was calibrated using the automatic calibration feature with the BTS. For each sample spot, an automatic acquisition with 240 laser shots was performed.

Mass Spectra Analysis

All protein spectra were identified by using the BDAL Bruker database (n = 8,468 MSPs), containing at least 3,000 different bacterial and fungi species, through the MBT Compass Explorer interface (v.4.1). The software attributed a log score value between 0 and 3.00. A score between 0 and 1.69 was considered as a not reliable identification. A score between 1.70 and 1.99 was considered as probable genus identification and scores from 2.00 to 2.29 as reliable genus identification and a probable species identification. Finally, a score between 2.30 and 3.00 was deemed as highly probable species identification.

Then, spectra were uploaded on FlexAnalysis v3.0 (Bruker Daltonics GmbH) and an internal calibration was carried out on the 4,365 m/z peak, identified as a 50 S ribosomal protein L36 by Zautner et al. (2016) in *Campylobacter*, which is shared by all samples and the BTS. Mass spectra were converted into mzML files and imported into BioNumerics v7.6 software platform (BioMérieux, Craponne, France). Spectra were pre-processed using the workflow described by Penny and collaborators [binned baseline (size = 77), Kaiser Window (size = 33), Moving bar (width = 129)], with a sound-to-noise ratio threshold of 10 (Penny et al., 2016). The peak detection parameters were the following: Continuous wavelet transformation (CWT) ridges, double peaks, and a relative intensity of 2%. Biological replicate spectra were summarized to create an average spectrum, or Main Spectra Profile (MSP), per isolate and extraction. Finally, a peak matching was performed on MSPs, resulting in 91 peaks.

Machine Learning Analysis

Pre-processing

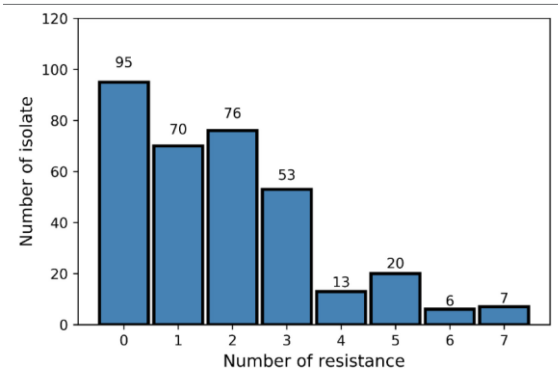
Tables including intensity values of the peak matching MSPs for the three types of extraction were exported into csv files (**Supplementary File 2**) for ML analysis using Python programming language (v3.7.6) and Scikit-learn package

TABLE 1 |

(A) Antimicrobial susceptibility patterns of *Campylobacter* isolates used in the present study.

		Resistant isolates	
Antibiotic classes	**Antibiotics**	***C. jejuni* (*n* = 224)**	***C. coli* (*n* = 116)**
	*Susceptible (S)**	70 (31.2%)	25 (21.6%)
Fluroquinolones	*Ciprofloxacin (Cip)*	123 (54.9%)	60 (51.7%)
Macrolides	*Erythromycin (Ery)*	2 (0.9%)	31 (26.7%)
Tetracyclines	*Tetracycline (Tet)*	90 (40.2%)	70 (60.3%)
Aminoglycosides	*Gentamycin (Gent)*	1 (0.4%)	11 (9.5%)
	Kanamycin (Kana)	18 (8.0%)	18 (15.5%)
	Streptomycin (Strep)	11 (4.9%)	35 (30.2%)
Beta-Lactams	*Ampicillin (Amp)*	90 (40.2%)	58 (50.0%)

(B) Diversity of antimicrobial resistance pattern in the collection.



Susceptible to all tested antimicrobials.

(v0.22.1) in Jupyter NoteBook (v6.0.3). Then, MSPs were grouped by their AMR profiles and eight distinct files have been created according their AMR classes and susceptibility, i.e., S, Cip^R, Tet^R, Amp^R, Ery^R, $Gent^R$, $Strep^R$, and $Kana^R$ (**Figure 1**). Category names (e.g., S and R) were binarized, where 0 and 1 represented MSPs susceptible and resistant to the AMR class studied, respectively. All peaks, here called features, were transformed using a Min-Max scaler which transformed values into the (0,1) range. Such a step is necessary to bring different variables at the same level, as variables that are measured at different scales may not contribute equally to the model fitting.

Feature's Selection

Dataset with many features, which could be redundant or irrelevant, may lead to an overcomplicated algorithm with low prediction accuracy and long training time. Feature selection is the process of choosing relevant features, to use in a classification model construction, either to improve accuracy scores or to boost performance. For this purpose, a meta-transformer based on a Random Forest estimator, implemented into scikit-learn library, was used to discard irrelevant features.

Model Selection

MSPs were randomly split into 70% training and 30% test datasets, with a stratification based on their binarized AMR profiles. The training dataset is implemented to build up a prediction model, while the test dataset is used as an external validation step of the trained model. For each studied AMR classes, Random Forests (RF), Logistic Regression (LR), and Naïve Bayes (NB) models were built, as they are common algorithms used in microbiology (Goodswen et al., 2021). RF is currently among the most used ML methods due to its robustness. It is essentially a collection of independent decision trees, where each tree could be different from the others, as the algorithm will make completely different random choices to make sure trees are distinct. Such algorithms make aggregated predictions using a group of decision trees. LR is a linear classifier, which predicts the probabilities of success and failure event. It is easy to implement and interpret and efficient to train. NB classifier assumes that the presence of a particular feature is not related to the presence of another feature. It is easy to interpret and is often applied for many medical applications. The area under the precision recall curve (AUPRC) was investigated to determine the most performant model (data not shown).

Tuning

Upon selection of the best performing model, it was optimized by looking for the best combination of hyper-parameters according to the F1-score, described in the metrics section. Hyper-parameters for each selected model were tuned by using an instance which generates candidates from a grid of given

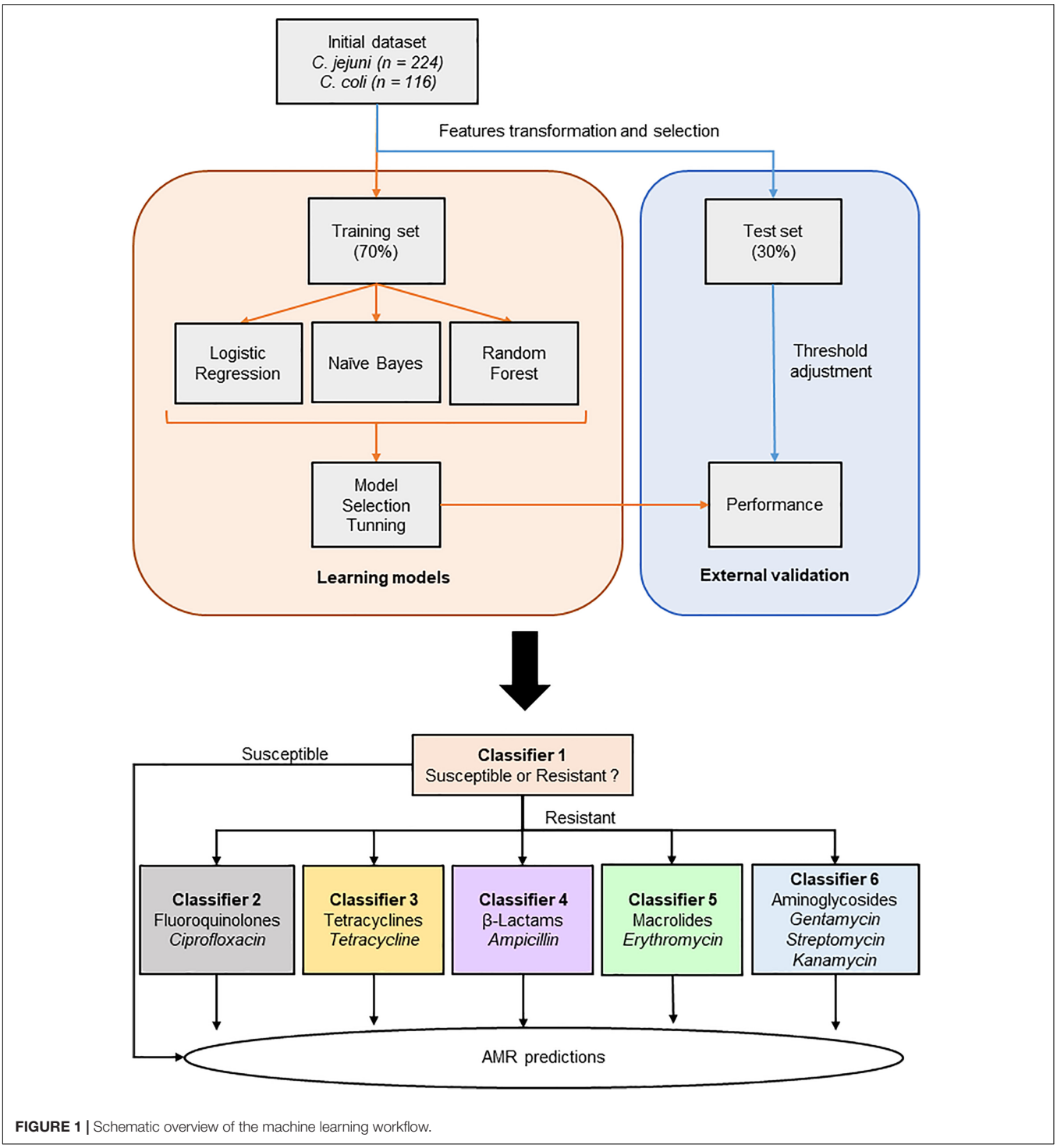


FIGURE 1 | Schematic overview of the machine learning workflow.

parameter values, a grid search, with a 10-fold cross validation, with a scoring method looking for the more optimized F1-score. K-fold cross validation is a resampling method, which estimates the performance of the ML model.

The 0.5 default probability score threshold may not represent an optimal interpretation and can result in poor performance. Therefore, a threshold adjustment was investigated to bring a higher predictive performance (Weis et al., 2020a). A threshold selection, for each classifier, based on their precision recall curve (PRC) was applied, according to the best F1-score. In the case of imbalance classes, like the current dataset, PRC can suggest an optimal threshold (Saito and Rehmsmeier, 2015). In this study, detection of resistant isolates (true positives) is the key point of the study. PRC is based on true positive values,

i.e., true positive and positive predictive values, among positive prediction. Hence, PRC relies on positive classes regardless of true negative value, making it a tool of choice for the study threshold selection. In the end, values less than the custom threshold are assigned to class 0, or susceptible, while value greater than or equal to the custom threshold are assigned to class 1, or resistant.

Performance and Metrics

As a next step, performance of the selected classifier needed to be assessed on data not yet seen by the model. For this, an external validation has been carried out by using the test dataset. Classification of spectra was summarized in a confusion matrix. From it, several performance metrics, such as the specificity, the recall, the precision or the positive predictive value (PPV), the negative predictive value (NPV), and area under the receiver operating characteristic curve (AUROC) and PRC were calculated. The PPV tells us how much we can trust the model when a resistant result is predicted, and in the other way, the NPV tells us how much we can trust the model when a sensitive result is predicted. The recall, also called sensitivity, measures how the model can find all positive units. The specificity refers to the model's ability to give a negative result when an isolate is susceptible. The ROC curve is a graphical way to represent the performance of the classifier for all threshold classifications, with the false-positive rate and true-positive rate as axis. Therefore, the AUROC can be used to measure the model's discriminative ability. Usually, an AUC of 0.5 is assimilated to a non-discriminative model, while 0.7–0.8 is considered acceptable, 0.8–0.9 is excellent, and more than 0.9 is considered outstanding (Hosmer et al., 2013). Along the same line, the PRC is a graphical visualization that combines the precision and the recall. The higher curve on the *y*-axis, the better the performance. Therefore, the AUPRC returns a value between 0 and 1, where 0 is the worst and 1 is the best. Finally, the F1-score is calculated from the precision and the recall. It conveys balance between the precision (PPV) and the recall (sensitivity).

Detailed information on ML analysis is shown in **Supplementary File 3**.

Biomarker Identification

Features of importance, based on RF algorithm trained on the whole dataset, were investigated to potentially identify already known antimicrobial resistance mechanisms or new antimicrobial targets. It rates how important each feature is for the decision tree. A score based on between 0 and 1 for each feature is calculated, where 0 means "Not used" and 1 highlighted a "perfect biomarker." Score for features of importance is computed as the mean and standard deviation of accumulation of the impurity decrease within each tree. Therefore, it describes the relevancy of a peak and, hence, can help to understand the biological problem. The five first features with the higher importance were checked in on Uniprot[1] according their mass in Da. Average theoretical masses were calculated using the online Expasy portal tool[2] based on Uniprot amino acid sequence.

Statistical Analysis

Effects of extraction methods on AMR predictions were analyzed based on analysis of variance (ANOVA) of the sum of AUPRCs of the different antimicrobial classifiers. ANOVA assumptions were verified with a Shapiro-Wilks and Levene tests. Shapiro-Wilks test determines if your data are normally distributed. The Leven test evaluates the equality of the variance. Differences were considered significant at $p < 0.05$.

RESULTS

Spectra Quality and Reproducibility

A total of 9,180 mass spectra were generated. An average identification log score of 2.0 was obtained for all spectra. Outlines, flatlines, and spectra not identified at the *Campylobacter* genus level were discarded for the analysis, resulting into 9,173 spectra. The latter was transformed into 1,020 MSPs, including 672 and 348 MSPs for *C. jejuni* and *C. coli,* respectively. Three different types of extractions, i.e., off-plate ethanol/acetonitrile extraction, direct deposit, and on-plate acid formic extraction, were carried out for both species. Hence, reproducibility was tested for the three biological replicates. Average similarities in percentage between the type of extraction and species are provided in **Figure 2**. For both species, no significant differences were observed between off- and on-plate extractions. Average similarity of means ranged from 77.1 to 92.7% between biological replicates for *C. jejuni* and *C. coli*, respectively.

Antimicrobial-Specific Screening

As a first step, different ML models, i.e., RF, LR, and NB, were trained for specific antimicrobials from different classes, regardless of the species identification to evaluate the potential of fast AM-screening without knowing the microbial identification. For this purpose, 1,020 MSPs, combining the three types of extractions and the two species, were split into a training and a validation set. The training set served to build the model, and the test set, to evaluate the performance of the model. Seven classifiers were built with RF and one with an LR algorithm. ROC and PR curves were computed to investigate the model's performance for each antibiotic (**Figure 3**), as well as other evaluation metrics such as sensitivity, specificity, PPV, and NPV summarized in **Table 2**.

Among the eight antimicrobials tested, three models performed better than the other considering both AUROC and AUPR curves. The best-performing model was the classifier allowing the distinction between resistant and completely susceptible isolates, with an area of 0.80 and 0.89 under the ROC and PR curves, respectively. The ciprofloxacin and tetracycline classifiers were the two other performant models according to their AUROC and the AUPR curves,

[1]https://www.uniprot.org/

[2]http://web.expasy.org/compute_pi/

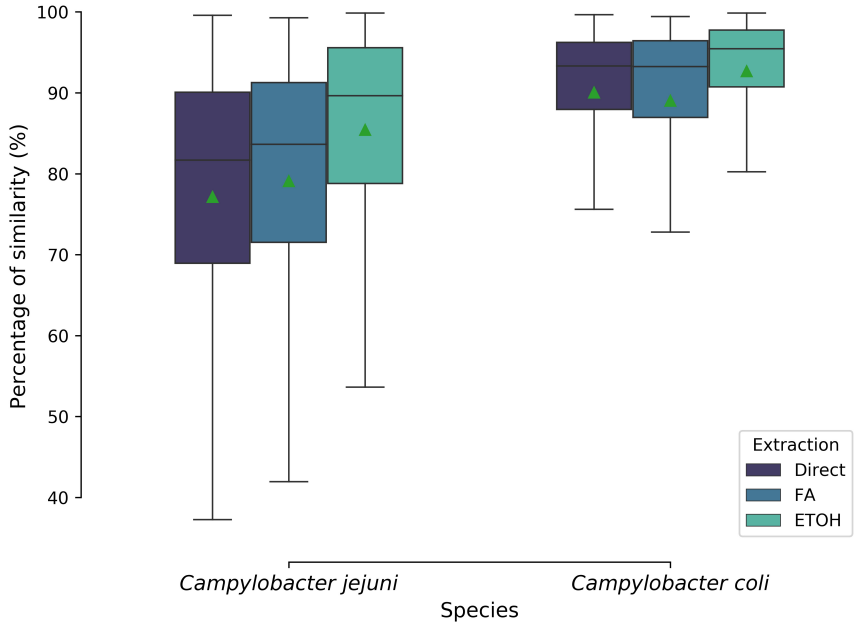


FIGURE 2 | Biological reproducibility of MALDI-TOF mass spectra based on their protein extraction type and species level. Boxplots show the isolates average similarities in percentage. Green triangle represented the mean. Direct, direct deposit; FA, formic acid extraction; EtOH, ethanol/acetonitrile extraction.

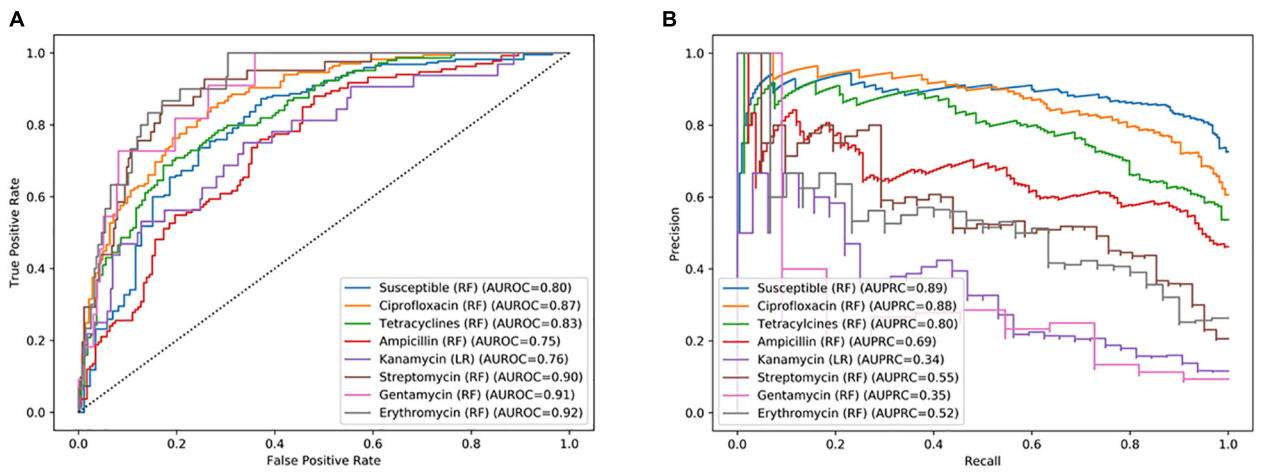


FIGURE 3 | (A) Receiver operating characteristic (ROC) curve and **(B)** recall–precision (PR) curves, and their related area under the curve, of specific antimicrobials based on combined *C. jejuni* and *C. coli* MALDI-TOF main protein spectra profiles (MSPs) of the test set (30%, *n* = 306). RF, Random Forest algorithm; LR, Logistic Regression algorithm; AUROC, Area Under the ROC Curve; AUPRC, Area Under the Precision Recall Curve.

an area of 0.87, 0.83, and 0.88, 0.80 under the AUROC and AUPRC, respectively (**Figure 3**). While the specificity was low for the three models, with a maximum of 63.8%, a sensitivity ranging from 87.5 and 92.3% was obtained (**Table 2**). Additionally, 74.6 and 85.7% of predicted values of the ciprofloxacin classifier could be reliable for resistant and susceptible values, respectively.

Remaining models had an AUROC of up to 0.92. However, considering the precision and the recall, they performed poorly. Indeed, the AUPR curve was between 0.34 and 0.69. Sensitivity and specificity may be high, but PPVs were low, e.g., 80.0, 88.4, and 42.8%, respectively, for the erythromycin model.

Species-Specific Screening

In a second phase, *C. coli* and *C. jejuni* MSPs were investigated separately to look over potential differences between tested antimicrobials. Previously, ROC and PR curves and their respective area under the curve have been computed, based on 202 and 105 MSPs, for the *C. jejuni* and *C. coli* test sets, respectively (**Figure 4**). As well, performance metrics were

TABLE 2 | Performance of retained machine learning classifier using combined *C. jejuni* and *C. coli* MALDI-TOF main protein spectra profiles (MSPs) of the test set (30%, *n* = 306 MSPs), grouped by the resistance profile.

Species	Antibiotics	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
C. jejuni and *C. coli* (*n* = 306 MSPs)	Susceptible* (*n* = 86)	92.3	45.3	81.2	69.6
	Ciprofloxacin (*n* = 165)	90.9	63.8	74.6	85.7
	Erythromycin (*n* = 30)	80.0	88.4	42.8	97.6
	Tetracycline (*n* = 144)	87.5	62.3	67.4	84.9
	Ampicillin (*n* = 133)	90.2	47.4	56.9	86.3
	Kanamycin (*n* = 32)	43.8	91.6	37.8	93.3
	Streptomycin (*n* = 41)	78.0	87.2	48.5	96.3
	Gentamycin (*n* = 11)	72.7	93.6	29.6	98.9

*Threshold applied for metrics calculation is based on the best F1-scores. PPV, positive predictive value; NPV, negative predictive value. *Susceptible to all tested antimicrobials.*

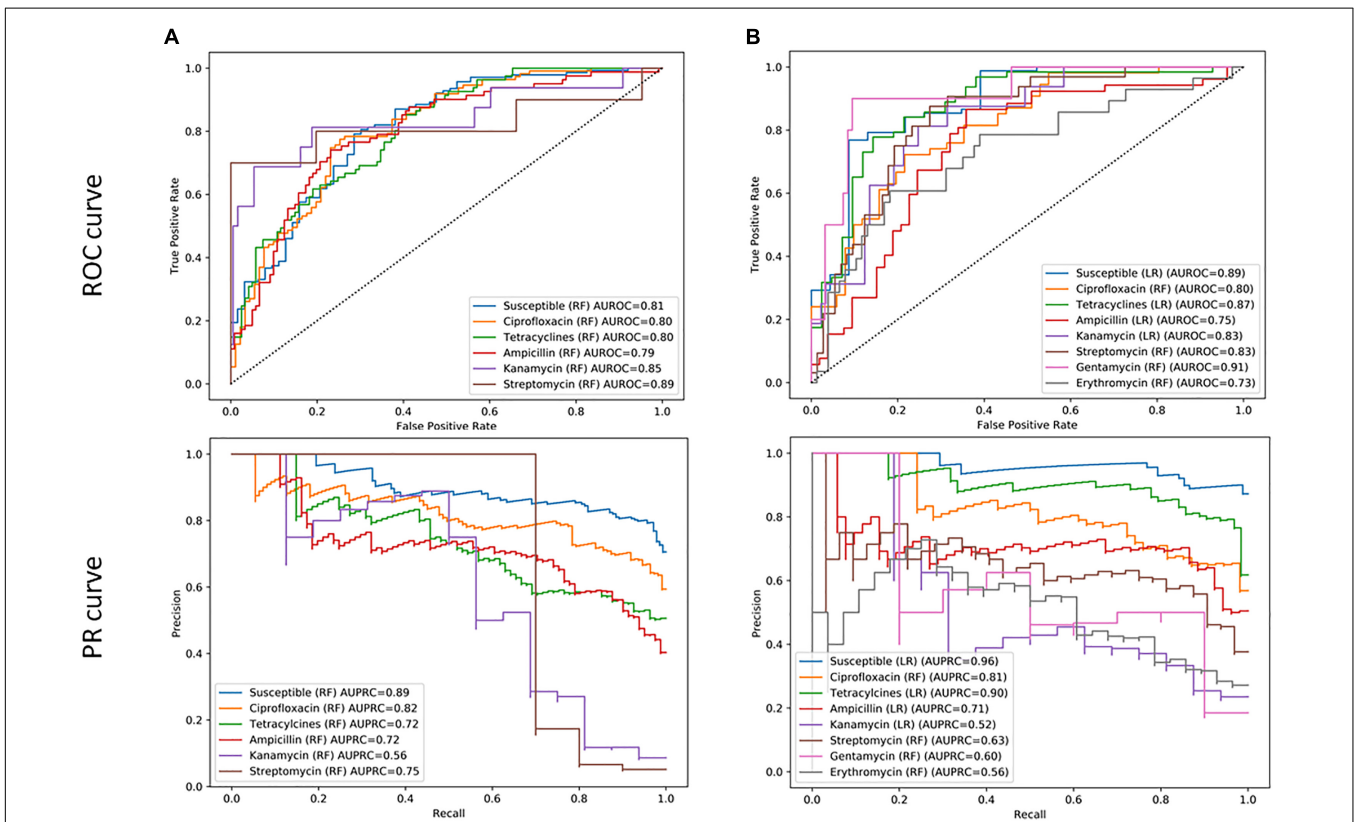


FIGURE 4 | Receiver operating characteristic (ROC) curve and recall–precision (PR) curves, and their related area under the curve, of specific antimicrobials based on 202 *C. jejuni* **(A)** and 105 *C. coli* **(B)** MALDI-TOF main protein spectra profiles (MSPs) of the test set (30%). RF, Random Forest algorithm; LR, Logistic regression algorithm; AUROC, Area under the ROC curve; AUPRC, Area under the precision–recall curve.

calculated (**Table 3**). Due to few gentamycin- and erythromycin-resistant isolates for *C. jejuni* in the initial collection (one and two, respectively), no model was built for these two antibiotics. RF and LR were once again fitting the best data. All six *C. jejuni* models were based on RF algorithms. Four models were built using LR and the remaining four were built using RF algorithms for *C. coli*.

As described in the specific antimicrobial section, the susceptible, ciprofloxacin, and tetracycline classifiers were the three best-performing models in both species, with an AUROC and AURP curve ranging from 0.80 to 0.89 and from 0.72 to 0.96, respectively (**Figure 4**). The susceptible classifier was the more performant model in both *C. jejuni* and *C. coli*. Tetracycline classifier was the second more effective model for *C. coli*, with an AUROC of 0.87 and AUPRC of 0.90, while it was the ciprofloxacin classifier for *C. jejuni*, with an AUROC of 0.80 and AUPRC of 0.82. Overall, sensitivity values up to 98.8% were obtained for these models. High PPVs and NPVs were obtained for susceptible classifiers. *C. coli* tetracycline classifier also performed well with a 79.2 and 92.9%

TABLE 3 | Performance of retained machine learning classifier using *C. jejuni* (*n* = 202 MSPs) and *C. coli* (*n* = 105 MSPs) MALDI-TOF main protein spectra profiles (MSPs) of the test set (30%), grouped by the resistance profile.

Species	Antibiotics	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
C. jejuni (*n* = 202 MSPs)	Susceptible* (*n* = 63)	92.8	55.6	82.2	77.8
	Ciprofloxacin (*n* = 111)	96.4	41.8	66.9	90.5
	Erythromycin (*n* = 2)	NA	NA	NA	NA
	Tetracycline (*n* = 81)	92.6	47.1	53.9	90.5
	Ampicillin (*n* = 81)	77.7	70.3	63.6	82.5
	Kanamycin (*n* = 16)	62.5	97.9	71.4	96.8
	Streptomycin (*n* = 10)	70.0	100.0	100.0	98.5
	Gentamycin (*n* = 1)	NA	NA	NA	NA
C. coli (*n* = 105 MSPs)	Susceptible* (*n* = 23)	98.8	60.9	90.0	93.3
	Ciprofloxacin (*n* = 54)	98.2	45.1	65.4	95.8
	Erythromycin (*n* = 28)	71.4	70.1	46.5	87.1
	Tetracycline (*n* = 63)	96.8	61.9	79.2	92.9
	Ampicillin (*n* = 52)	86.5	64.1	70.3	82.9
	Kanamycin (*n* = 16)	62.5	86.5	45.5	92.7
	Streptomycin (*n* = 32)	84.3	75.3	60.0	91.7
	Gentamycin (*n* = 10)	70.0	93.7	53.8	96.7

*Threshold applied for metrics calculation is based on the best F1-scores. PPV, positive predictive value; NPV, negative predictive value. *Susceptible to all tested antimicrobials. NA, Not applicable due to few isolates in the category.*

for PPV and NPV, respectively. Surprisingly, the ciprofloxacin classifier was less efficient in both species. Indeed, a lower PPV was obtained, i.e., 10% differences, in comparison with previous results where the microbial identification was not taken into consideration. For erythromycin, kanamycin, and gentamycin classifiers, observations described in the previous section could be assessed.

Differences were observed for the ampicillin and streptomycin classifier for *C. coli* and *C. jejuni*. *C. jejuni* streptomycin's classifier performed more efficiently than the one of *C. coli*. PPVs and NPVs of 100 and 98.5%, against 60.0 and 91.7%, were calculated, respectively. *C. coli* ampicillin's classifier was more performant than that of *C. jejuni*, while similar AUROC and AUPR curves were found. Indeed, PPVs and NPVs of 70.3 and 82.9% against 63.6 and 82.5% were calculated for *C. coli* and *C. jejuni*, respectively (**Table 3**).

Protein Extraction Impact on Resistance Predictions

Thirdly, methods of extraction, i.e., direct deposit, FA on-plate, and EtOH/ACN off-plate extraction, were investigated to check potential variation for specific antimicrobials. Thereby, MSPs acquired for each extraction for both *C. jejuni* (n = 224 MSPs) and *C. coli* (n = 116 MSPs) were used to build a specific ML model per antimicrobial. Models are compared in **Figure 5**. The ANOVA resulted in 0.976 and 0.936 ($p > 0.05$) values for *C. jejuni* and *C. coli,* respectively. Therefore, the null hypothesis, i.e., there is no difference between extraction methods, is retained.

Nevertheless, in the case of the *C. coli* gentamycin's classifier, while the performance is low for the EtOH/ACN extraction (AUPRC = 0.23), the classifier for the direct deposit is more efficient (AUPRC = 0.92). Features of extractions for both classifiers were investigated. For the EtOH/ACN classifier, 2,356.29 Da was the more important feature. For the direct deposit classifier, 10,323.79 Da was the more important feature. While these features in a model were particularly important, they were the less important features in the other model. The 10,323.79 Da peak was detected in both extractions, while softly shifting for the EtOH/ACN, i.e., 10,333.67 Da. The 2,356.29 peak was not detected in the direct deposit (**Figure 6**).

Biomarkers: Antimicrobial Resistance Mechanisms

RF classifiers performing the best, i.e., susceptible, ciprofloxacin, and tetracycline, while microbial species is not known, were used to retrieve features of importance. Then, the Uniprot database was investigated to potentially identify each feature according their mass in Dalton, regardless post-translational modifications. **Table 4** summarizes the top five features for each classifier. When several proteins had the same mass, proteins with the most probable function linked to AMR were retained. No protein for *C. jejuni* or *C. coli* was identified at 6,436.22 Da. The DNA methyltransferase at 6,436 Da was in *Helicobacter pylori*, a closely related genus of *Campylobacter*.

DISCUSSION

Several reports described MALDI-TOF MS as a more time- and cost-effective alternative approach to current classic AST methods (Hrabák et al., 2013; Oviaño and Bou, 2019). Being combined with ML, such an approach may be even more relevant for AST in routine diagnostics (Weis et al., 2020b). However, to our knowledge, no study implying relevant foodborne pathogens for AMR screening has been published yet. Therefore, the scope of this study was to consider whether a mass spectrometry

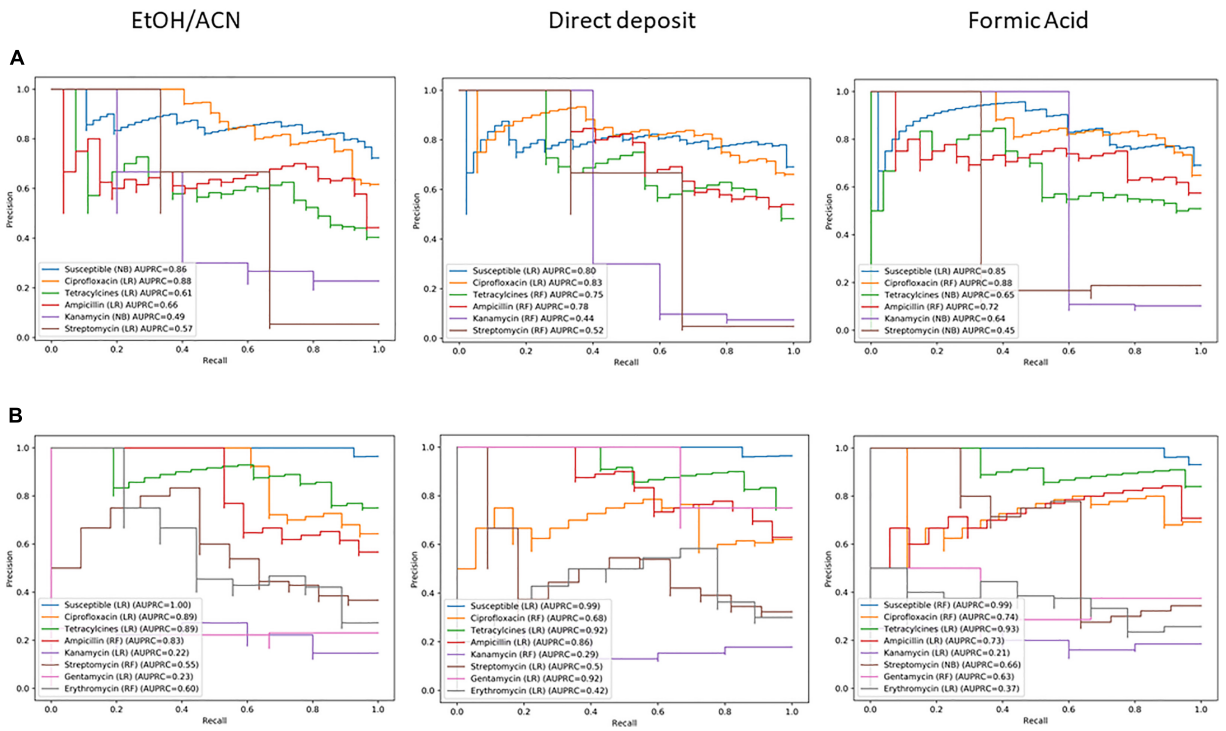


FIGURE 5 | Comparison of precision-recall curves for the three-extraction tested on **(A)** *C. jejuni* (*n* = 68 MSPs) and **(B)** *C. coli* (*n* = 35 MSPs) of the test set (30%). EtOH/ACN: complete ethanol/acetonitrile-based proteins extraction. RF, Random Forest algorithm; LR, Logistic Regression algorithm; NB, Nave Bayes algorithm.

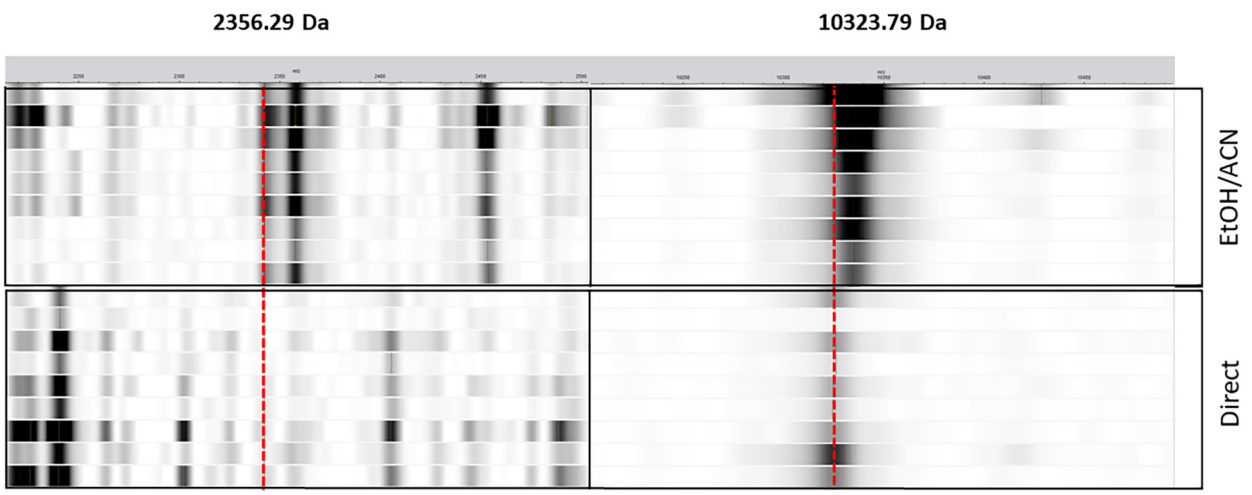


FIGURE 6 | Pseudogel view representation of mass spectra from *C. coli* from the direct deposit (direct, *n* = 8) and the ethanol/acetonitrile off-plate extraction (EtOH/ACN, *n* = 8). The *x*-axis represents the mass-to-charge (m/z) ratio in Da. Strips intensities is function of the peak intensity. The red dashed lines represent the observed peaks, i.e., 2,356.29 and 10,323.79 Da.

technique combined with an ML approach could be utilized for a combined rapid species identification and AMR screening for foodborne pathogens.

The main result of this study was to observe whether mass spectra with 91 protein peaks selected by automatic peak-matching could predict with a high average sensitivity and precision the strains' susceptibility and resistance to ciprofloxacin and tetracycline, independent of the microbial species identification. Therefore, these models were missing very few resistant isolates. Similarly, Weis and colleagues, computed an AUROC for 42 different antibiotics on a large "real-world" clinical dataset by combining multiple species

TABLE 4 | Top five ranking of Random Forest features of importance.

Classifier	Rank	Features (Da)	Average theoretical mass (Da)	Protein	UniProt ID
Susceptible	1	8460.76	8460.07	Transcriptional regulator	A0A1T1ZLP8
	2	3257.41	3256.98	GNAT family N-acetyltransferase	A0A6N3Q833
	3	5867.81	5867.86	ATP-binding protein	A0A2A5MAC7
	4	2766.98	2767.13	Poly(A) polymerase	A0A5T1K937
	5	4365.25	4364.39	50 S ribosomal protein L36	A0A1E7P1M9
Ciprofloxacin	1	6436.22	6435.55	DNA methyltransferase*	A0A438RVN3*
	2	2766.98	2767.13	Poly(A) polymerase	A0A5T1K937
	3	2241.84	2241.67	Type II toxin-antitoxin system HicB family antitoxin	A0A691V648
	4	3257.41	3256.98	GNAT family N-acetyltransferase	A0A6N3Q833
	5	7083.30	7083.03	MmgE/PrpD family protein	A0A4Y8C2R1
Tetracycline	1	4365.25	4364.39	50 S ribosomal protein L36	A0A1E7P1M9
	2	2766.98	2767.13	Poly(A) polymerase	A0A5T1K937
	3	7083.30	7083.03	MmgE/PrpD family protein	A0A4Y8C2R1
	4	6436.22	6435.55	DNA methyltransferase*	A0A438RVN3*
	5	2713.95	2713.06	Superoxide dismutase	A0A431FY74

*Da, Dalton. *Identified in the closely related genus Helicobacter pylori (former Campylobacter pylori).*

(Weis et al., 2020a). They pointed out that they reached AUROC values above 0.90 for 23 of the tested antibiotics. Such results support the idea that mass spectra could provide far more than simple species information. Nevertheless, in the literature, most of the publications focused on specific species such as *S. aureus*, *Escherichia coli*, and *Klebsiella pneumoniae*. Additionally, they mainly analyzed one type of antimicrobial classes, e.g., glycopeptides such as vancomycin (Mather et al., 2016; Asakura et al., 2018; Wang et al., 2018; Candela et al., 2021). For example, Asakura et al. (2018) obtained a sensitivity of 99.0% and a specificity of 88.0% while comparing vancomycin-susceptible and heterogeneous vancomycin intermediately resistant *S. aureus*. Wang et al. (2018) obtained similar results with a 77.0 and 81.4% sensitivity and specificity, respectively, for the same comparison. When comparing *C. jejuni* and *C. coli* separately and for different antimicrobials, we found that susceptible, ciprofloxacin, and tetracycline classifiers were the three best-performing models in both species, while the others performed less accurately. Similarly to other studies, a sensitivity ranging from 92.6 to 98.8% was obtained for both species and the three performant classifiers. Weis et al. (2020a) also looked at species-specific antimicrobial resistance prediction for *S. aureus*, *E. coli*, and *K. pneumoniae*. They reported an AUROC ranging from 0.77 to 0.81, and an AUPRC ranging from 0.52 to 0.70 for ciprofloxacin predictions. In the current study, similar AUROC values were found but a higher AUPRC was observed with 0.82 and 0.81 for *C. jejuni* and *C. coli*, respectively, meaning that the current model may accurately predict ciprofloxacin-resistant isolates. Considered as a critically important antimicrobial, ciprofloxacin is widely used for the treatment of broad human bacterial infections, including enteric ones (WHO, 2019). Therefore, early screening of its resistance may play an essential role for the administration of the definitive antimicrobial therapy. Nevertheless, the comparison between the different studies is intricate to perform due to the number of isolates, the genus analyzed, the type of extraction, as well as the type of algorithm used. In the current study, classifiers performing poorly, i.e., kanamycin, streptomycin, gentamycin, and erythromycin, were subject to a highly imbalanced dataset, with an average of 10/90 resistant/susceptible ratio, instead of a close 50/50 ratio one (e.g., 36 gentamycin-resistant MSPs for 984 gentamycin-susceptible MSPs). Precision disparities were observed for the ciprofloxacin, ampicillin, and streptomycin classifiers of both species, in comparison to classifiers not considering the species level. While such differences could be attributed to the unbalanced number of resistant isolates for ampicillin and streptomycin, the ciprofloxacin classifier was in contrast well balanced. The ciprofloxacin classifier may be less effective for predictions, while looking specifically at the species level. In the end, prediction based on protein mass spectra grouped by AMR, regardless of bacterial species, may be the best option for an efficient and swift AMR-screening. Such observations might also be explained by average similarity differences obtained between *C. jejuni* and *C. coli*. Cuénod and Egli (2021), Cuénod et al. (2021) reported that the preparation protocol used, the duration of incubation, maintenance of the device, for example, could potentially impact the quality of the spectra. Inevitably it may have influenced the final prediction for both species. Hypothetically, such observations may also show that AMR screening by MALDI-TOF MS is going beyond the bacterial genus or species and might be directly linked to the resistance mechanism and protein/metabolite expression itself. To our knowledge, this is the first study establishing that ML and MALDI-TOF MS could be applied for AMR screening of foodborne pathogens, such as *Campylobacter* spp.

Nevertheless, in the current study, the specificity was not as high as the specificity described by the previously mentioned studies. While creating the ML pipeline, sensitivity was chosen as the most important parameter to adjust the threshold score during the tuning part. Hence, the optimal threshold was selected based on the F1-score, meaning the

best compromise between higher sensitivity and precision, specific to each classifier. Classifiers guiding antibiotic therapy decision must have high sensitivity (Weis et al., 2020a). On the one hand, assuming an isolate is susceptible, while it is resistant, may lead to an ineffective treatment and eventually have an important impact on patient management. On the other hand, assuming an isolate is resistant, while it is susceptible, may still lead to an effective treatment. However, while seeking and picking to have high sensitivity, it will inevitably decrease the specificity, by decreasing it. In the previously cited reports, threshold adjustments were not mentioned. Therefore, threshold adjustment may be a key step while elaborating ML pipeline for routine laboratories based on MALDI-TOF mass spectra.

The impact of protein extraction methods was also evaluated. Indeed, the EtOH/ACN extraction is the most popular extraction protocol when it comes to research investigations. However, the direct deposit and the on-plate FA extractions are the most straightforward methods used in routine laboratories. No significant differences were observed between the direct deposit, the FA on-plate, and the EtOH/ACN extraction. Therefore, in order to rapidly obtain straightforward AMR assessment information, the application of the direct deposit method could be applied for species identification as well as AMR screening in *Campylobacter*. Interestingly, *C. coli* gentamycin classifier performance was different between EtOH/ACN extraction and the direct deposit. Indeed, with a simple biological smear on the MALDI-TOF target, gentamycin's prediction was more precise. Surprisingly, the absence of the 2,356.29 Da peak resulted in a higher AUPRC for the direct deposit classifier. In the literature, the loss of a specific peak between different types have already been described (Josten et al., 2014). However, in their case, the loss of a protein happened during the ethanol washing step of the EtOH/ACN extraction. Thus, the peptide was only present during a direct deposit measurement. However, to confirm our observation, additional gentamycin-resistant isolates should be analyzed as currently too few gentamycin isolates are present in the current dataset.

Along the same line, putative biomarkers have been identified for each class of studied antibiotics by looking into RF algorithm features of importance. Majority of these proteins, such as transcriptional regulator, ATP-binding, GCN5-related N-acetyltransferase, DNA-methyltransferase, toxin-antitoxin system, PrpD, and superoxide dismutase proteins had a direct or indirect link with already known antibiotic resistance, tolerance, or spread mechanisms in different genera of bacteria (e.g., *Salmonella*, *Enterococcus*, *Escherichia*, *Mycobacterium*, and *Pseudomonas*) (Draker and Wright, 2004; Yugendran and Harish, 2016; Hicks et al., 2018; Kang et al., 2018; Martins et al., 2018; Su et al., 2018; Shaheen et al., 2020). Nevertheless, *Campylobacter*'s AMR mechanisms are either chromosomal mutations, such as the single mutation C257T in the *gyrA* gene or the A207G mutation in the 23 S rRNA gene for ciprofloxacin and erythromycin, respectively, or acquired genes, such as *tet(O)*, bla_{OXA}*-61* and *aph(3')-III* for tetracycline, ampicillin, and gentamycin resistances, respectively (Payot et al., 2006; Iovine, 2013). Overall, these mechanisms are working in synergy with the *cmeABC* efflux pump or porines, such the Major-Out-Membrane Porines (MOMP) (Lin et al., 2002). Over the biomarkers identified as relevant by RF susceptible classifier, the GCN5-related N-acetyltransferase and the 50 S ribosomal protein L36 may be linked to already known aminoglycosides or tetracyclines resistance mechanisms of *Campylobacter*, respectively. On one hand, aminoglycoside-modifying enzymes, such as acetyltransferase [e.g., *aac(6′)-Ie–aph(2″)-If2]* were already detected in gentamycin-resistant *Campylobacter* isolates (Zhao et al., 2016). On the other hand, the Tet(O) ribosomal protection protein is known to bind on both 30S and 50S subunits, conferring tetracycline resistance (Li et al., 2013). Interestingly, the L36 proteins were the first feature of importance highlighted for the tetracycline classifier. Identification of specific proteins directly implied to AMR mechanisms, while using MALDI-TOF MS within the 2–20 kDa range, could be problematic (Welker and Van Belkum, 2019). Indeed, proteins responsible for resistances are large proteins (e.g., GyrA = 96,974 Da). Therefore, in case an indicative biomarker is identified, it may not be a necessary protein conferring the resistance itself, but it may be a protein or peptide co-coded on the plasmid of the protein responsible of the resistance (Lau et al., 2014). Therefore, the 4,365.25 m/z peak may be a biosignature linked to the presence of the *tet(O)* gene. In the literature, two protein biomarkers, i.e., 3,665.79 *m/z* and 6,036.59 *m/z*, have been reported to be a potential biomarker of the tetracycline resistance in other bacterial genera (Sabença et al., 2020; Sousa et al., 2020). However, these biomarkers were not observed here. Along the same line, the 6,436.22 Da protein was considered as the most important feature for the ciprofloxacin's classifier. The protein was identified as a DNA methylase in *H. pylori*, formerly related to the *Campylobacter* genus. Yugendran and Harish put in light the hypothesis that ciprofloxacin-resistance in *E. coli* may be induced by DNA methylation, leading to the possible involvement of some mechanism other than the quinolone-resistance determining region (QRDR) capable of inducing fluoroquinolone resistance (Yugendran and Harish, 2016). While the single point mutation in *gyrA* represents the major fluoroquinolones resistance mechanism in *Campylobacter*, such venue may be worth exploring in the future. Other potential ciprofloxacin biomarkers, neighboring 6,300 Da, were put recently in light for other *E. coli* (Sousa et al., 2020) and *Enterococcus* (Sabença et al., 2020; Sousa et al., 2020). Nevertheless, interpretation on the biological role of features may be cautiously interpreted, and a peptide sequencing by tandem mass spectrometry should be performed to assess the real biological function of these biomarkers.

Little is known on the impact of such approaches as described here on the health management potential cost savings in clinical practice. Weis and colleagues affirmed in their study that the application of such workflow provided a treatment guidance 12–72 h earlier than classical approaches and to have a significant impact on the physician–patient workflow (Weis et al., 2020a). It is worth mentioning that the ML is intended for supporting the decision making process. Therefore, it is a

support giving guidance on possible resistance outcomes that lead early antibiotherapy in a specific direction. ML may be used as an AMR screening tool, displaying an alert message on the MALDI-TOF MS microbial identification report, when the isolate is classified as a positive category value. It is already the case for several Bruker subtyping modules (e.g., MRSA, cfiA positive or bla$_{KPC}$ modules). Therefore, instead of giving an empirical treatment until the AMR confirmation by reference AST, the patient's antibiotherapy may be defined faster (e.g., 24 h earlier).

Phenotypic antibiogram should still follow up to establish the AMR profile and, in case, reorient the antibiotherapy. Additionally, 2025 AMR monitoring of food-producing isolates, such as ESBL/AmpC/carbapenemase-producing *E. coli*, will be done by WGS (Aerts et al., 2019). Therefore, a combination of MALDI-TOF MS, ML, and WGS could be an interesting monitoring tool with a relevant impact on the control of the emergence of AMR in the European Union. As well, the application of MALDI-TOF MS in microbiology for lipid investigation has conceptualized several breakthroughs for AMR screening (Bruker, 2019; Furniss et al., 2019; Dortet et al., 2020). In case of the ability of such method to distinct microbial lipids directly from body fluids such as serum, blood, and urine, there will be no need of a culture step (Solntceva et al., 2021). So far, only the last-line treatment for multidrug-resistant Gram-negative bacteria, i.e., polymyxin, has been investigated without a ML approach. Lipidomics combined to artificial intelligence may be a new venue to explore AMR problem cases that proteomics could not solve. However, there is still a stony way before the long-term implementation of ML in routine laboratories for AMR screening. Nevertheless, a single protein mass spectra may be used in the future as an utmost "One-fits all" diagnostics tool for: species identification, AMR screening, and genetic diversity (Feucherolles et al., 2021).

Several limitations of our study are offered for consideration. First, the employed dataset might be considered as relatively small to train an ML algorithm properly. Indeed, lack of data could lead a model to overfit or underfit the data. Several models (e.g., gentamycin or kanamycin) were trained on heavy unbalanced classes, which is not recommended to build a robust and reliable tool for AMR predictions. Therefore, extra isolates resistant to these antimicrobials should be added to the current dataset. Additionally, only three ML algorithms, i.e., RF, LR, and NB, were tested. The support vector machine algorithm was not included in the study, while it is also a widely used algorithm for AMR predictions. Another limitation of the study is the use of disk-diffusion antibiograms, which—while being a valid and highly reproducible method to characterize an isolate as resistant or susceptible—do not allow quantifying the minimal inhibitory concentration (MIC) of a given antibiotic. Additionally, it would have been possible to test for further antibiotics, e.g., carbapenems. The final limitation of this study could be the fact that the RF model, used for putative biomarkers identification, was trained on the whole dataset. Indeed, under these settings, there is no proof that these biomarkers could work in a given analysis. For such investigations, the model should have been trained on a split dataset, including a training and test set, with a 70/30% ratio, respectively.

CONCLUSION

On the one hand, MALDI-TOF MS in combination with supervised ML may be a powerful tool for the fast screening of foodborne pathogens such as *C. coli* and *C. jejuni*, which might be susceptible, ciprofloxacin, or tetracycline resistant. On the other hand, other antimicrobials tested, i.e., ampicillin, gentamycin, kanamycin, streptomycin, and erythromycin, did not provide good results to reach a conclusion for its application under clinical settings, due to unbalance datasets. Nonetheless, this work could serve as a proof-of-concept, and future research should include other important foodborne pathogens such as *Salmonella* spp. Our approach has the potential to obtain the following information from one single protein spectrum analysis: species identification, antimicrobial susceptibility patterns, and genetic diversity.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

MF carried out MALDI-TOF MS, machine learning, and data analysis, and drafted the manuscript with MN and CR. MN and CR isolated and performed the identification and AMR characterization of the *Campylobacter* collection. SB and DM supplied extra *Campylobacter* strains from their respective clinical laboratories. CR, SB, and DM provided their expert critical point of view on the current work. SL gave access to his lab for all mass spectrometry analysis. H-MC supervised the project. CP wrote the project proposal and obtained funding. All authors contributed to the formal analysis, writing, review, and editing of the manuscript, read and agreed to the published version of the manuscript.

FUNDING

This research was supported by the Luxembourg National Research Fund (FNR): MICROH-DTU FNR PRIDE program (No. 11823097).

ACKNOWLEDGMENTS

We acknowledge Katleen Vranckx from BioMérieux for her help in enabling straightforward pre-processing of the data analysis with BioNumerics. We warmly thank Nathalie Geoders

from the Luxembourg Institute of Science (LIST) and technology and Aurélien Savart from the company Lemonads, Luxembourg, for their machine learning expertise support. We acknowledge as well Monique Perrin and Marie Meo from the Laboratoire National de Santé, Luxembourg for selecting multi-drugs resistant isolates from their national AMR surveillance activities. We would also like to thank Fatù Djabi from the Laboratoire National de Santé, Luxembourg, and Dominique Claude from the Laboratoire de Médecine Vétérinaire de l'Etat, Luxembourg, for their technical support throughout the project. We are also thankful to Louise Hock from the LIST for critical reading as layman in the field.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.804484/full#supplementary-material

Supplementary File 1 | Details of the *Campylobacter* spp. collection used for the study.

Supplementary File 2 | *C. jejuni* and C. coli peak matching table for the EtOH/ACN, FA extraction, and direct deposit.

Supplementary File 3 | Example of one of the supervised machine learning Python workflows.

REFERENCES

Aerts, M., Battisti, A., Hendriksen, R., Kempf, I., Teale, C., Tenhagen, B. A., et al. (2019). Technical specifications on harmonised monitoring of antimicrobial resistance in zoonotic and indicator bacteria from food-producing animals and food. *EFSA J.* 17:5709. doi: 10.2903/j.efsa.2019.5709

Arena, F., Giani, T., Pollini, S., Viaggi, B., Pecile, P., and Rossolini, G. M. (2017). Molecular antibiogram in diagnostic clinical microbiology: advantages and challenges. *Future Microbiol.* 12, 361–364. doi: 10.2217/fmb-2017-0019

Asakura, K., Azechi, T., Sasano, H., Matsui, H., Hanaki, H., Miyazaki, M., et al. (2018). Rapid and easy detection of low-level resistance to vancomycin in methicillin-resistant *Staphylococcus aureus* by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *PLoS One* 13:e0194212. doi: 10.1371/journal.pone.0194212

Barlam, T. F., Cosgrove, S. E., Abbo, L. M., Macdougall, C., Schuetz, A. N., Septimus, E. J., et al. (2016). Implementing an antibiotic stewardship program: guidelines by the infectious diseases society of America and the society for healthcare epidemiology of America. *Clin. Infect. Dis.* 62, e51–e77. doi: 10.1093/cid/ciw118

Benkova, M., Soukup, O., and Marek, J. (2020). Antimicrobial susceptibility testing: currently used methods and devices and the near future in clinical practice. *J. Appl. Microbiol.* 129, 806–822. doi: 10.1111/jam.14704

Bookstaver, P. B., Nimmich, E. B., Smith, T. J., Justo, J. A., Kohn, J., Hammer, K. L., et al. (2017). Cumulative effect of an antimicrobial stewardship and rapid diagnostic testing bundle on early streamlining of antimicrobial therapy in Gram-negative bloodstream infections. *Antimicrob. Agents Chemother.* 61, 1–10. doi: 10.1128/AAC.00189-17

Bortolaia, V., Kaas, R. S., Ruppe, E., Roberts, M. C., Schwarz, S., Cattoir, V., et al. (2020). ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* 75, 3491–3500. doi: 10.1093/jac/dkaa345

Boucheikhchoukh, M., Laroche, M., Aouadi, A., Dib, L., Benakhla, A., Raoult, D., et al. (2018). MALDI-TOF MS identification of ticks of domestic and wild animals in algeria and molecular detection of associated microorganisms. *Comput. Immunol. Microbiol. Infect. Dis.* 57, 39–49. doi: 10.1016/j.cimid.2018.05.002

Bredtmann, C. M., Krücken, J., Murugaiyan, J., Kuzmina, T., and von Samson-Himmelstjerna, G. (2017). Nematode species identification—current status, challenges and future perspectives for cyathostomins. *Front. Cell Infect. Microbiol.* 7:1–8. doi: 10.3389/fcimb.2017.00283

Bruker (2019). *Bruker Launches MALDI Biotyper Sirius at ASM Microbe Conference*. Available online at: www.asm.org (accessed July 8, 2019)

Caballé, N. C., Castillo-Sequera, J. L., Gómez-Pulido, J. A., Gómez-Pulido, J. M., and Polo-Luque, M. L. (2020). Machine learning applied to diagnosis of human diseases: a systematic review. *Appl. Sci.* 10, 1–27. doi: 10.3390/app10155135

Candela, A., Arroyo, M. J., Mancera, L., Microbiology, C., General, H., Software, C. B., et al. (2021). Rapid and reproducible MALDI-TOF-based method for detection vancomycin- resistant *Enterococcus faecium* using classifying algorithms. *bioRxiv* [preprint]. doi: 10.1101/2021.06.23.449689

Chlebicz, A., and Śliżewska, K. (2018). Campylobacteriosis, salmonellosis, yersiniosis, and listeriosis as zoonotic foodborne diseases: a review. *Int. J. Environ. Res. Public Health* 15, 1–29. doi: 10.3390/ijerph15050863

Clark, A. E., Kaleta, E. J., Arora, A., and Wolk, D. M. (2013). Matrix-assisted laser desorption ionization-time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. *Clin. Microbiol. Rev.* 26, 547–603. doi: 10.1128/CMR.00072-12

Cuénod, A., and Egli, A. (2021). "Advanced applicationsof maldi-tof ms—typing and beyond," in *Application and Integration of Omics-powered Diagnostics in Clinical and Public Health Microbiology*, eds J. Moran-Gilad and Y. Yagel (Heidelebrg/Berlin, Germany: Springer), 153–173. doi: 10.1007/978-3-030-62155-1_9

Cuénod, A., Foucault, F., Pflüger, V., and Egli, A. (2021). Factors associated with MALDI-TOF mass spectral quality of species identification in clinical routine diagnostics. *Front Cell Infect Microbiol.* 16:1–15. doi: 10.3389/fcimb.2021.646648

Del Chierico, F., Di Cave, D., Accardi, C., Santoro, M., Masotti, A., D'Alfonso, R., et al. (2016). Identification and typing of free-living *Acanthamoeba* spp. by MALDI-TOF MS biotyper. *Exp. Parasitol.* 170, 82–89. doi: 10.1016/j.exppara.2016.09.007

Dortet, L., Bonnin, R. A., Le Hello, S., Fabre, L., Bonnet, R., Kostrzewa, M., et al. (2020). Detection of colistin resistance in *Salmonella enterica* using MALDIxin test on the routine MALDI biotyper sirius mass spectrometer. *Front. Microbiol.* 11:1–6. doi: 10.3389/fmicb.2020.01141

Draker, K. A., and Wright, G. D. (2004). Molecular mechanism of the enterococcal aminoglycoside 6' -N-acetyltransferase': role of GNAT-conserved residues in the chemistry of antibiotic inactivation. *Biochemistry* 43, 446–454. doi: 10.1021/bi035667n

EFSA, and ECDC (2021). The european union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2018/2019. *EFSA J.* 19, 1–179. doi: 10.2903/j.efsa.2021.6490

Emele, M. F., Karg, M., Hotzel, H., Graaf-van Bloois, L., Groß, U., Bader, O., et al. (2019a). Differentiation of campylobacter fetus subspecies by proteotyping. *Eur. J. Microbiol. Immunol.* 9, 62–71. doi: 10.1556/1886.2019.00006

Emele, M. F., Možina, S. S., Lugert, R., Bohne, W., Masanta, W. O., Riedel, T., et al. (2019b). Proteotyping as alternate typing method to differentiate campylobacter coli clades. *Sci. Rep.* 9, 1–11. doi: 10.1038/s41598-019-40842-w

Fabre, A., Oleastro, M., Nunes, A., Santos, A., Sifré, E., Ducournau, A., et al. (2018). Whole-genome sequence analysis of multidrug-resistant campylobacter isolates: a focus on aminoglycoside resistance determinants. *J. Clin. Microbiol.* 56, 1–12. doi: 10.1128/jcm.00390-18

Fall, B., Lo, C. I., Samb-Ba, B., Perrot, N., Diawara, S., Gueye, M. W., et al. (2015). The ongoing revolution of MALDI-TOF mass spectrometry for microbiology reaches tropical Africa. *Am. J. Trop. Med. Hyg.* 92, 641–647. doi: 10.4269/ajtmh.14-0406

Feucherolles, M., Cauchie, H., and Penny, C. (2019a). MALDI-TOF mass spectrometry and specific biomarkers: potential new key for swift identification of antimicrobial resistance in foodborne pathogens. *Microorganisms* 7, 1–16. doi: 10.3390/microorganisms7120593

Feucherolles, M., Poppert, S., Utzinger, J., and Becker, S. L. (2019b). MALDI-TOF mass spectrometry as a diagnostic tool in human and veterinary helminthology: a systematic review. *Parasit. Vectors* 12, 1–13. doi: 10.1186/s13071-019-3493-9

Feucherolles, M., Nennig, M., Becker, S. L., Martiny, D., Losch, S., Penny, C., et al. (2021). Investigation of MALDI - TOF mass spectrometry for assessing the molecular diversity of *Campylobacter jejuni* and comparison with MLST and cgMLST: a luxembourg one - health study. *Diagnostics* 11, 1–17. doi: 10.3390/diagnostics11111949

Florio, W., Tavanti, A., Ghelardi, E., and Lupetti, A. (2018). MALDI-TOF MS applications to the detection of antifungal resistance: state of the art and future perspectives. *Front. Microbiol.* 9:1–7. doi: 10.3389/fmicb.2018.02577

Furniss, R. C. D., Dortet, L., Bolland, W., Drews, O., Sparbier, K., Bonnin, R. A., et al. (2019). Detection of colistin resistance in *Escherichia coli* by use of the MALDI biotyper sirius mass spectrometry system. *J. Clin. Microbiol.* 57, 1–7. doi: 10.1128/JCM.01427-19

Ge, M. C., Kuo, A. J., Liu, K. L., Wen, Y. H., Chia, J. H., Chang, P. Y., et al. (2017). Routine identification of microorganisms by matrix-assisted laser desorption ionization time-of-flight mass spectrometry: success rate, economic analysis, and clinical outcome. *J. Microbiol. Immunol. Infect.* 50, 662–668. doi: 10.1016/j.jmii.2016.06.002

Gibson, D., and Costello, C. E. (2000). Mass spectrometry of biomolecules. *Separation Sci. Technol.* 2, 299–327. doi: 10.1016/S0149-6395(00)80053-7

Goodswen, S. J., Barratt, J. L. N., Kennedy, P. J., Kaufer, A., Calarco, L., and Ellis, J. T. (2021). Machine learning and applications in microbiology. *FEMS Microbiol. Rev.* 015, 1–19. doi: 10.1093/femsre/fuab015

Hicks, N. D., Yang, J., Zhang, X., Zhao, B., Grad, Y. H., Liu, L., et al. (2018). Clinically prevalent mutations in *Mycobacterium tuberculosis* alter propionate metabolism and mediate multidrug tolerance. *Nat. Microbiol.* 3, 1032–1042. doi: 10.1038/s41564-018-0218-3

Hosmer, D. W., Lemeshow, S., and Sturdivant, R. X. (2013). "Assessing the fit of the model," in *Applied Logistic Regression*, (John Willey & Sons, Inc.), 177. doi: 10.1002/9781118548387

Hrabák, J., Chudác ková, E., and Walková, R. (2013). Matrix-assisted laser desorption ionization-time of flight (MALDITOF) mass spectrometry for detection of antibiotic resistance mechanisms: from research to routine diagnosis. *Clin. Microbiol. Rev.* 26, 103–114. doi: 10.1128/CMR.00058-12

Iles, R. K., Zmuidinaite, R., Iles, J. K., Carnell, G., Sampson, A., and Heeney, J. L. (2020). A clinical MALDI-ToF mass spectrometry assay for SARS-CoV-2: rational design and multi-disciplinary team work. *Diagnostics* 10, 1–15. doi: 10.1101/2020.08.22.20176669

Iovine, N. M. (2013). Resistance mechanisms in campylobacter jejuni. *Virulence* 4, 230–240. doi: 10.4161/viru.23753

Josten, M., Dischinger, J., Szekat, C., Reif, M., Al-Sabti, N., Sahl, H. G., et al. (2014). Identification of agr-positive methicillin-resistant *Staphylococcus aureus* harbouring the class a mec complex by MALDI-TOF mass spectrometry. *Int. J. Med. Microbiol.* 304, 1018–1023. doi: 10.1016/J.IJMM.2014.07.005

Kang, S. M., Kim, D. H., Jin, C., and Lee, B. J. (2018). A systematic overview of type II and III toxin-antitoxin systems with a focus on druggability. *Toxins* 10, 1–21. doi: 10.3390/toxins10120515

Lau, A. F., Wang, H., Weingarten, R. A., Drake, S. K., Suffredini, A. F., Garfield, M. K., et al. (2014). 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.* 52, 2804–2812. doi: 10.1128/JCM.00694-14

Li, W., Atkinson, G. C., Thakor, N. S., Allas, U., Lu, C. C., Yan Chan, K., et al. (2013). Mechanism of tetracycline resistance by ribosomal protection protein Tet(O). *Nat. Commun.* 4, 1–8. doi: 10.1038/ncomms2470

Lin, J., Overbye Michel, L., and Zhang, Q. (2002). CmeABC functions as a multidrug efflux system in campylobacter jejuni. *Antimicrob. Agents Chemother.* 46, 2124–2131. doi: 10.1128/AAC.46.7.2124-2131.2002

Mangioni, D., Viaggi, B., Giani, T., Arena, F., D'Arienzo, S., Forni, S., et al. (2019). Diagnostic stewardship for sepsis: the need for risk stratification to triage patients for fast microbiology workflows. *Future Microbiol.* 14, 169–174. doi: 10.2217/fmb-2018-0329

Martins, D., McKay, G., Sampathkumar, G., Khakimova, M., English, A. M., and Nguyen, D. (2018). Superoxide dismutase activity confers (p)ppgppmediated antibiotic tolerance to stationary-phase *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.* 115, 9797–9802. doi: 10.1073/pnas.1804525115

Mather, C. A., Werth, B. J., Sivagnanam, S., SenGupta, D. J., and Butler-Wu, S. M. (2016). Rapid detection of vancomycin-intermediate *Staphylococcus aureus* by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J. Clin. Microbiol.* 54, 883–890. doi: 10.1128/JCM.02428-15

Mitchell, S. L., and Alby, K. (2017). Performance of microbial identification by MALDI-TOF MS and susceptibility testing by VITEK 2 from positive blood cultures after minimal incubation on solid media. *Eur. J. Clin. Microbiol. Infect. Dis.* 36, 2201–2206. doi: 10.1007/s10096-017-3046-0

Nami, Y., Imeni, N., and Panahi, B. (2021). Application of machine learning in bacteriophage research. *BMC Microbiol.* 21:1–8. doi: 10.1186/s12866-021-02256-5

Olkkola, S., Juntunen, P., Heiska, H., Hyytiäinen, H., and Hänninen, M. L. (2010). Mutations in the rpsL gene are involved in streptomycin resistance in campylobacter coli. *Microb. Drug Resist.* 16, 105–110. doi: 10.1089/mdr.2009.0128

On, S. L. W., Zhang, J., Cornelius, A. J., and Anderson, T. P. (2021). Markers for discriminating campylobacter concisus genomospecies using MALDI-TOF analysis. *Curr. Res. Microb. Sci.* 2, 1–5. doi: 10.1016/j.crmicr.2020.100019

Oviaño, M., and Bou, G. (2019). Matrix-assisted laser desorption ionization–time of flight mass spectrometry for the rapid detection of antimicrobial resistance mechanisms and beyond. *Clin. Microbiol. Rev.* 32, 1–29. doi: 10.1128/CMR.00037-18

Payot, S., Bolla, J. M., Corcoran, D., Fanning, S., Mégraud, F., and Zhang, Q. (2006). Mechanisms of fluoroquinolone and macrolide resistance in campylobacter spp. *Microbes Infect.* 8, 1967–1971. doi: 10.1016/j.micinf.2005.12.032

Penny, C., Grothendick, B., Zhang, L., Borror, C. M., Barbano, D., Cornelius, A. J., et al. (2016). A designed experiments approach to optimizing MALDI-TOF MS spectrum processing parameters enhances detection of antibiotic resistance in *Campylobacter jejuni*. *Front. Microbiol.* 7:1–9. doi: 10.3389/fmicb.2016.00818

Robert, M. G., Cornet, M., Hennebique, A., Rasamoelina, T., Caspar, Y., Pondérand, L., et al. (2021). Maldi-tof ms in a medical mycology laboratory: on stage and backstage. *Microorganisms* 9, 1–16. doi: 10.3390/microorganisms9061283

Rodriguez-Granger, J., Mendiola, J. C., and Martínez, L. A. (2018). "Identification of mycobacteria by matrix-assisted laser desorption ionization–time-of-flight mass spectrometry," in *The Use of Mass Spectrometry Technology (MALDI-TOF) in Clinical Microbiology*, ed. F. Cobo (Academic Press), 181–195. doi: 10.1016/B978-0-12-814451-0.00013-7

Rodríguez-Sánchez, B., Cercenado, E., Coste, A. T., and Greub, G. (2019). Review of the impact of MALDI-TOF MS in public health and hospital hygiene, 2018. *Euro Surveill* 24, 1–12. doi: 10.2807/1560-7917.ES.2019.24.4.1800193

Rotcheewaphan, S., Lemon, J. K., Desai, U. U., Henderson, C. M., and Zelazny, A. M. (2019). Rapid one-step protein extraction method for the identification of mycobacteria using MALDI-TOF MS. *Diagn. Microbiol. Infect. Dis.* 94, 355–360. doi: 10.1016/j.diagmicrobio.2019.03.004

Roth, S., Berger, F. K., Link, A., Nimmesgern, A., Lepper, P. M., Murawski, N., et al. (2021). Application and clinical impact of the RESIST-4 O.K.N.V. rapid diagnostic test for carbapenemase detection in blood cultures and clinical samples. *Eur. J. Clin. Microbiol. Infect. Dis.* 40, 423–428. doi: 10.1007/s10096-020-04021-4

Rybicka, M., Miłosz, E., and Bielawski, K. P. (2021). Superiority of MALDI-TOF mass spectrometry over real-time PCR for SARS-CoV-2 RNA detection. *Viruses* 13, 1–11. doi: 10.3390/v13050730

Sabença, C., de Sousa, T., Oliveira, S., Viala, D., Théron, L., Chambon, C., et al. (2020). Next-generation sequencing and maldi mass spectrometry in the study of multiresistant processed meat vancomycin-resistant enterococci (VRE). *Biology (Basel).* 9, 1–17. doi: 10.3390/biology9050089

Saito, T., and Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* 10:1–21. doi: 10.1371/journal.pone.0118432

Shaheen, A., Tariq, A., Shehzad, A., Iqbal, M., Mirza, O., Maslov, D. A., et al. (2020). Transcriptional regulation of drug resistance mechanisms in *Salmonella*: where we stand and what we need to know. *World J. Microbiol. Biotechnol.* 36, 1–8. doi: 10.1007/s11274-020-02862-x

Singhal, N., Kumar, M., Kanaujia, P. K., and Virdi, J. S. (2015). MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front. Microbiol.* 6:1–16. doi: 10.3389/fmicb.2015.00791

Sogawa, K., Watanabe, M., Ishige, T., Segawa, S., Miyabe, A., Murata, S., et al. (2017). Rapid discrimination between methicillin-sensitive and

methicillin-resistant *Staphylococcus aureus* using MALDI-TOF mass spectrometry. *Biocontrol Sci.* 22, 163–169. doi: 10.4265/bio.22.163

Solntceva, V., Kostrzewa, M., and Larrouy-Maumus, G. (2021). Detection of species-specific lipids by routine MALDI TOF mass spectrometry to unlock the challenges of microbial identification and antimicrobial susceptibility testing. *Front. Cell Infect. Microbiol.* 10:1–12. doi: 10.3389/fcimb.2020.621452

Sousa, T., De, Viala, D., Th, L., and Chambon, C. (2020). Putative protein biomarkers of *Escherichia coli* antibiotic multiresistance identified by MALDI mass spectrometry. *Biology (Basel)* 9, 1–14. doi: 10.3390/biology9030056

Su, W., Kumar, V., Ding, Y., Ero, R., Serra, A., Lee, B. S. T., et al. (2018). Ribosome protection by antibiotic resistance ATP-binding cassette protein. *Proc. Natl. Acad. Sci. U.S.A.* 115, 5157–5162. doi: 10.1073/pnas.1803313115

Sy, I., Margardt, L., Ngbede, E. O., Adah, M. I., Yusuf, S. T., Keiser, J., et al. (2021). Identification of adult *Fasciola* spp. using matrix-assisted laser/desorption ionization time-of-flight (maldi-tof) mass spectrometry. *Microorganisms* 9, 1–11. doi: 10.3390/microorganisms9010082

Tahir, D., Almeras, L., Varloud, M., Raoult, D., Davoust, B., and Parola, P. (2017). Assessment of MALDI-TOF mass spectrometry for filariae detection in *Aedes aegypti* mosquitoes. *PLoS Negl. Trop. Dis.* 11:1–18. doi: 10.1371/journal.pntd.0006093

Tandina, F., Laroche, M., Davoust, B., Doumbo, K., and Parola, P. (2018). Blood meal identification in the cryptic species *Anopheles gambiae* and *Anopheles coluzzii* using MALDI-TOF MS. *Parasite* 25, 1–7. doi: 10.1051/parasite/2018041

Tang, W., Ranganathan, N., Shahrezaei, V., and Larrouy-Maumus, G. (2019). MALDI-TOF mass spectrometry on intact bacteria combined with a refined analysis framework allows accurate classification of MSSA and MRSA. *PLoS One* 14:1–16. doi: 10.1371/journal.pone.0218951

Wang, H. Y., Chen, C. H., Lee, T. Y., Horng, J. T., Liu, T. P., Tseng, Y. J., et al. (2018). Rapid detection of heterogeneous vancomycin-intermediate *Staphylococcus aureus* based on matrix-assisted laser desorption ionization time-of-flight: using a machine learning approach and unbiased validation. *Front. Microbiol.* 9:1–10. doi: 10.3389/fmicb.2018.02393

Weis, C. V., Cuénod, A., Rieck, B., Llinares-López, F., Dubuis, O., Graf, S., et al. (2020a). Direct antimicrobial resistance prediction from MALDI-TOF mass spectra profile in clinical isolates through machine learning. *bioRxiv* 1, 1–35. doi: 10.1101/2020.07.30.228411

Weis, C. V., Jutzeler, C. R., and Borgwardt, K. (2020b). Machine learning for microbial identification and antimicrobial susceptibility testing on MALDI-TOF mass spectra: a systematic review. *Clin. Microbiol. Infect.* 26, 1310–1317. doi: 10.1016/j.cmi.2020.03.014

Welker, M., and Van Belkum, A. (2019). One system for all: is mass spectrometry a future alternative for conventional antibiotic susceptibility testing? *Front. Microbiol.* 10:1–12. doi: 10.3389/fmicb.2019.02711

Wendel, T., Feucherolles, M., Rehner, J., Keiser, J., Poppert, S., Utzinger, J., et al. (2021). Evaluating different storage media for identification of *Taenia saginata* proglottids using MALDI-TOF mass spectrometry. *Microorganisms* 9:2006. doi: 10.3390/microorganisms9102006

WHO (2019). *Critically Important Antimicrobials for Human Medicine 6th Re.Vision 2018. Ranking of Medically Important Antimicrobials for Risk Management of Antimicrobial Resistance Due to Non-Human Use*. Available online at: https://apps.who.int/iris/bitstream/handle/10665/312266/9789241515528-eng.pdf?ua=1 (accessed May 15, 2019).

Yoon, E., and Jeong, S. H. (2021). MALDI-TOF mass spectrometry technology as a tool for the rapid diagnosis of antimicrobial resistance in bacteria. *Antibiotics* 10, 1–13. doi: 10.3390/antibiotics10080982

Yugendran, T., and Harish, B. N. (2016). Global DNA methylation level among ciprofloxacin-resistant clinical isolates of *Escherichia coli*. *J. Clin. Diagnostic Res.* 10, 27–30. doi: 10.7860/JCDR/2016/19034.7830

Zautner, A. E., Lugert, R., Masanta, W. O., Weig, M., Groß, U., and Bader, O. (2016). Subtyping of *Campylobacter jejuni* ssp. doylei isolates using mass spectrometry-based phyloproteomics (MSPP). *J. Visual. Exp. JoVE* 116, 1–9. doi: 10.3791/54165

Zhao, S., Tyson, G. H., Chen, Y., Li, C., Mukherjee, S., Young, S., et al. (2016). Whole-genome sequencing analysis accurately predicts antimicrobial resistance phenotypes in campylobacter spp. *Appl. Environ. Microbiol.* 82, 459–466. doi: 10.1128/AEM.02873-15

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 08 March 2022
doi: 10.3389/fmicb.2022.827451

Large-Scale Samples Based Rapid Detection of Ciprofloxacin Resistance in *Klebsiella pneumoniae* Using Machine Learning Methods

Chunxuan Wang[1,2], Zhuo Wang[1,3], Hsin-Yao Wang[4,5], Chia-Ru Chung[6], Jorng-Tzong Horng[6,7], Jang-Jih Lu[4,8,9] and Tzong-Yi Lee[1,2]**

[1] Warshel Institute for Computational Biology, The Chinese University of Hong Kong, Shenzhen, China, [2] School of Data Science, The Chinese University of Hong Kong, Shenzhen, China, [3] School of Life Sciences, University of Science and Technology of China, Hefei, China, [4] Department of Laboratory Medicine, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan, [5] Ph.D. Program in Biomedical Engineering, Chang Gung University, Taoyuan City, Taiwan, [6] Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan, [7] Department of Bioinformatics and Medical Engineering, Asia University, Taichung City, Taiwan, [8] Department of Medical Biotechnology and Laboratory Science, Chang Gung University, Taoyuan, Taiwan, [9] Department of Medicine, College of Medicine, Chang Gung University, Taoyuan, Taiwan

OPEN ACCESS

Edited by:
Karsten Becker,
University Medicine Greifswald, Germany

Reviewed by:
Dana Marshall,
Meharry Medical College, United States
Miriam Cordovana,
Bruker Daltonik GmbH, Germany

***Correspondence:**
Jang-Jih Lu
janglu45@gmail.com
Tzong-Yi Lee
leetzongyi@cuhk.edu.cn

Specialty section:
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

Received: 02 December 2021
Accepted: 17 January 2022
Published: 08 March 2022

Citation:
Wang C, Wang Z, Wang H-Y, Chung C-R, Horng J-T, Lu J-J and Lee T-Y (2022) Large-Scale Samples Based Rapid Detection of Ciprofloxacin Resistance in Klebsiella pneumoniae Using Machine Learning Methods. Front. Microbiol. 13:827451. doi: 10.3389/fmicb.2022.827451

Klebsiella pneumoniae is one of the most common causes of hospital- and community-acquired pneumoniae. Resistance to the extensively used quinolone antibiotic, such as ciprofloxacin, has increased in *Klebsiella pneumoniae*, which leads to the increase in the risk of initial antibiotic selection for *Klebsiella pneumoniae* treatment. Rapid and precise identification of ciprofloxacin-resistant *Klebsiella pneumoniae* (CIRKP) is essential for clinical therapy. Nowadays, matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) is another approach to discover antibiotic-resistant bacteria due to its shorter inspection time and lower cost than other current methods. Machine learning methods are introduced to assist in discovering significant biomarkers from MALDI-TOF MS data and construct prediction models for rapid antibiotic resistance identification. This study examined 16,997 samples taken from June 2013 to February 2018 as part of a longitudinal investigation done by Change Gung Memorial Hospitals (CGMH) at the Linkou branch. We applied traditional statistical approaches to identify significant biomarkers, and then a comparison was made between high-importance features in machine learning models and statistically selected features. Large-scale data guaranteed the statistical power of selected biomarkers. Besides, clustering analysis analyzed suspicious sub-strains to provide potential information about their influences on antibiotic resistance identification performance. For modeling, to simulate the real antibiotic resistance predicting challenges, we included basic information about patients and the types of specimen carriers into the model construction process and separated the training and testing sets by time. Final performance reached an area under the receiver operating characteristic curve (AUC) of 0.89 for support vector machine (SVM) and extreme gradient boosting (XGB) models. Also, logistic regression and random forest models both achieved AUC around 0.85. In conclusion, models provide sensitive forecasts of CIRKP, which may aid in early

antibiotic selection against *Klebsiella pneumoniae*. The suspicious sub-strains could affect the model performance. Further works could keep on searching for methods to improve both the model accuracy and stability.

Keywords: antibiotic susceptibility test, MALDI-TOF MS, machine learning, *Klebsiella pneumonia*, ciprofloxacin resistance

BACKGROUND AND INTRODUCTION

Klebsiella pneumoniae (*K. pneumoniae*) is one of the most common hospital- and community-acquired bacterial infections (Ashurst and Dawson, 2021). Extensively used quinolone antibiotics, which include ciprofloxacin, play a significant role in *K. pneumoniae* treatment. Increases in the proportion of ciprofloxacin-resistant *K. pneumoniae* (CIRKP) and the long inspection time of traditional antimicrobial susceptibility testing (AST) could lead to incorrect initial antibiotic treatment that will squander the essential treatment time of patients (Burckhardt and Zimmermann, 2018). Methods for rapid and precise identification of CIRKP are critical for clinical *K. pneumoniae* infection treatment. With the assistance of matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) technology, inspection time of AST, and strain typing of infectious bacteria could decrease to less than 2 h after cell culture (van Belkum et al., 2017; Florio et al., 2018). Introducing machine learning methods to antibiotic resistance identification could assist in further improving the inspection speed and lowering the cost (Weis et al., 2020) and discovering potential antibiotic markers from MALDI-TOF MS data. In this study, prediction models for CIRKP in the Taiwan area are constructed based on large-scale mass spectrum data and non-spectrometric information of patients.

Although *K. pneumoniae* is normally harmless and can be found in the human intestine, it can cause serious infections in other parts of the body, such as pneumoniae, urinary tract infection, sepsis, etc. Especially, multi-drug resistant and carbapenem-resistant *K. pneumoniae* (CRKP) has become a great threat to public health, whose overall 30-days mortality rate has been reported to be greater than 40% due to the limited antibiotic options after infection (Tumbarello et al., 2012). Thus, most recent studies related to *K. pneumoniae* focus on creating new identification methods and finding potential resistant biomarkers. However, as a broad-spectrum quinolone antibiotic, ciprofloxacin is one of the widely used antibiotics for the treatments of infection caused by *K. pneumoniae* in clinical therapy. The increasing reported CIRKP (Sanchez et al., 2013; Zhou et al., 2016) could also be a severe problem under clinical circumstances. Identification models for different bacterial species and antibiotics are important to validate the feasibility of the generalization of machine learning-based antibiotic resistance detection using MALDI-TOF MS data.

Clinical prescriptions usually heavily depend on the AST result to guide the initial antibiotic selection and avoid inefficient treatment. Traditional AST procedure usually requires 24 h for plate culture and an additional 24 h for the antimicrobial susceptibility testing (Kumar et al., 2009). The delay of efficient antibiotic treatment will increase the mortality rate, and especially if the patient is seriously infected. In recent years, the polymerase chain reaction (PCR) method has been applied to rapidly detect genes of *K. pneumoniae* related to quinolone resistance, such as mutations on type II and type IV DNA topoisomerase genes (Ruiz, 2003; Nordmann and Poirel, 2005). Besides, abnormal expression of outer cytomembrane efflux pump and plasmid-mediated resistance genes are also proved to be quinolone resistance mechanisms and can be detected by genomic and proteomic methods (Martínez-Martínez et al., 1998). However, restriction to the current gene library, high labor intensity, and excessive cost of the tests are still practical problems for generalizing genome tests. Nowadays, utilizing the MALDI-TOF MS could significantly shorten the testing time and lower the inspection cost. Achievements have been made in both identifying antibiotic-resistant bacteria from laboratory plate cultures and directly from the specimens of patients (Clark et al., 2013; Patel, 2015; Singhal et al., 2015; Angeletti, 2016; Arca-Suárez et al., 2017; Sandalakis et al., 2017; Tré-Hardy et al., 2017). In 2016, spectrum peak at 11.109 m/z was confirmed related to plasmid-mediated CRKP with gene (Gaibani et al., 2016). Besides, polypeptide at 3,043 m/z is proved to be a fragment of PBP2a, which participants in the methicillin resistance process of Staphylococcus aureus (MRSA) (Sogawa et al., 2017). Those study results demonstrate that MALDI-TOF MS can find specific mass peaks with potential biological meanings. It may also detect antibiotic resistance profiles of large protein-involved mechanisms in a low mass-to-charge ratio range. Thus, antibiotic-resistant biomarkers obtained from MALDI-TOF MS data may not only serve as evidence for bacterial type identification, but we may also even be able to find resistant bacterial strains with unrevealed mechanisms.

Predicting models constructed by machine learning methods has achieved high accuracy in identifying antibiotic bacterial strains. Taking the identification of MRSA as an example, the support vector machine (SVM) models show identification accuracy of around 90% (Sogawa et al., 2017; Tang et al., 2019), and the model in the study of Liu et al. (2021) shows the area under the receiver operating characteristic (ROC) curve (AUC) of 0.89 for SVM model and 0.87 for random forest model (RF). Especially, according to Wang et al. (2021), a logistic regression (LR) model is trained based on over 20,000 samples and independently validated by another data set with more than 5,000 samples, and finally achieves a predicting AUC of 0.85. Large samples used in their study and external validation highly improve the reliability of the machine learning model. However, most previous studies apply statistical analysis and construct models only based on a small number of samples (usually the total

samples are no more than 1,000), which could significantly limit the statistical power of the analysis and the reliability of models.

This study performed an antibiotic resistance analysis and modeling based on 16,697 samples collected from a longitudinal study from June 2013 to February 2018 and AST results were taken as references. Raw spectra data were first preprocessed by peak smoothing and baseline correction. After that, peaks with signal-to-noise ratio of 2 were selected for further analysis. Since *K. pneumoniae* may cause serious infections in various body parts, this study analyzed six common *K. pneumoniae* specimen carriers, which include blood samples (B), body fluid samples (F), wound samples (W), respiratory samples (R), urinary samples (U), and other types (O). Besides, to simulate the real clinical antibiotic resistance predicting problems, patients' basic information, including gender and age, were also included in model construction, coupling with separating training and testing sets by time. Statistical tests were applied to select significant features for modeling. Final performances achieved predicting AUC of around 0.89 for SVM and extreme gradient descent boosting (XGB) models and AUC 0.85 for LR and RF models. Moreover, this study also included a clustering analysis based on an unsupervised learning method to provide potential information of different *K. pneumoniae* sub-strains in the data set and quantified the influence of suspicious sub-strains on the model performances.

MATERIALS AND METHODS

The modeling procedure of this research is shown in the flow chart in **Figure 1**, which primarily contains five distinct components. The process began with specimen collection and mass spectral data processing, followed by data cleaning for quality control. After that, data processing aimed to unify the spectra's dimensions and generated dummy variables for categorical variables. Finally, the relevant variables selected in the feature selection step were used to create binary classification models with balancing methods.

Data Cleaning

The samples used in this study were collected from a longitudinal study on ciprofloxacin that ran from June 2013 to February 2018 done by the Change Gung Memorial Hospitals (CGMH) at the Linkou branch. The longitudinal research included specimens of 16,697 participants in total. To guarantee the data quality, 915 samples labeled as intermediate and 4 incomplete samples were eliminated. In total, 15,782 samples were selected for further analysis, of which 11,354 were ciprofloxacin-susceptible Klebsiella pneumoniae (CISKP) and 4,428 were CIRKP.

Specimen Preparation and Spectrum Preparation Methods

All the specimen samples were collected as daily routine examinations. Specimens with different carrier types were cultured with the most appropriate methods. Blood specimens were cultured in the trypticase soy broth (Becton Dickinson, MD) and an automated detection system (BD BACTECTM FX; Becton Dickinson) was utilized for detecting positive blood culture results. After that, blood from positive blood culture bottles was inoculated on blood plate (BP) agar for subculture. Fluid specimens and respiratory specimens were inoculated on BP agar, eosin methylene blue (EMB) agar, CNA agar, and chocolate agar. Additionally, some of the fluid specimens were cultured on thioglycolate broth. Specimens obtained from wounds were first rinsed with 1.2 mL of 0.9% saline solution and then inoculated on BP, EMB, CNA, and chocolate agars. For urine specimens, only BP and EMB agars were utilized for the plate culture. All the agars and broths used for specimen cultures were from Becton Dickinson and they were incubated into a 37°C CO_2 incubator for 18–24 h. MALDI-TOF MS was performed using selected single colonies. CIRKP specimens were identified from CISKP specimens through the disk diffusion method under the instruction of Clinical and Laboratory Standards Institute guideline M100 (CLSI M100). CISKP samples and CIRKP samples were determined by ATS breakpoints listed in CLSI M100 (Clinical and Laboratory Standards Institute [CLSI], 2018).

The selected colonies were analyzed by MALDI-TOF MS (Microflex LT MALDI-TOF System, Bruker Daltonik GmbH) following the operating instructions created by the manufacturer. First, cultivated colonies were smeared onto the MALDI steel target plate with the addition of formic acid (1 μL, 70%) and then dried at 25°C degrees. Then, a matrix solution (α-cyano-4-hydroxycinnamic acid, 100 mg/mL, 50% acetonitrile with 2.5% trifluoroacetic acid) was added to the spots, and the samples were dried at room temperature. Spectrum data with mass-to-charge ratio (m/z) between 2,000 and 20,000 were then collected using Microflex LT MALDI-TOF analyzer in a linear mode (accelerating voltage, 20 kV; nitrogen laser frequency: 60 Hz; 240 laser shott). The raw spectrum data were first calibrated with an external standard calibrator (Bruker Daltonics Bacterial Test Standard), and then peak smoothing (Savitzky-Golay filter) and baseline correction (Tof-hat filter) were applied. Finally, peaks with signal-to-noise ratio 2 were selected for further analysis.

Spectrometric Data and Non-spectrometric Data Processing

To unify the dimensions of spectrum data along with alleviating the influence of peak shift, spectral data was grouped into pseudo-ion vectors each with 900 pseudo-ions. Pseudo-ions were calculated by first grouping the mass-to-charge ratio into bins of width 20 m/z, and then selecting the peak with maximum intensity ratio. Using the pseudo-ion peak k of the sample i as an example:

$$P_{ik} = Max_j \left(R_{ij} \cdot 1_{\left\{ (2000+20(k-1)) \le M_{ij} < (2000+20k) \right\}} \right), \quad k \in [1, 900] \cap \boldsymbol{Z}$$

where M_{ij} and R_{ij} represent the mass-to-charge ratio and the intensity ratio of the j^{th} peak of the sample i, respectively, $1_{\{.\}}$ is the indicator function. After that, pseudo-ion vectors were row-bind together to form the spectral data matrix. $\overrightarrow{P_k}$ indicates the pseudo-ion k (the k^{th} column of the spectral matrix) and the

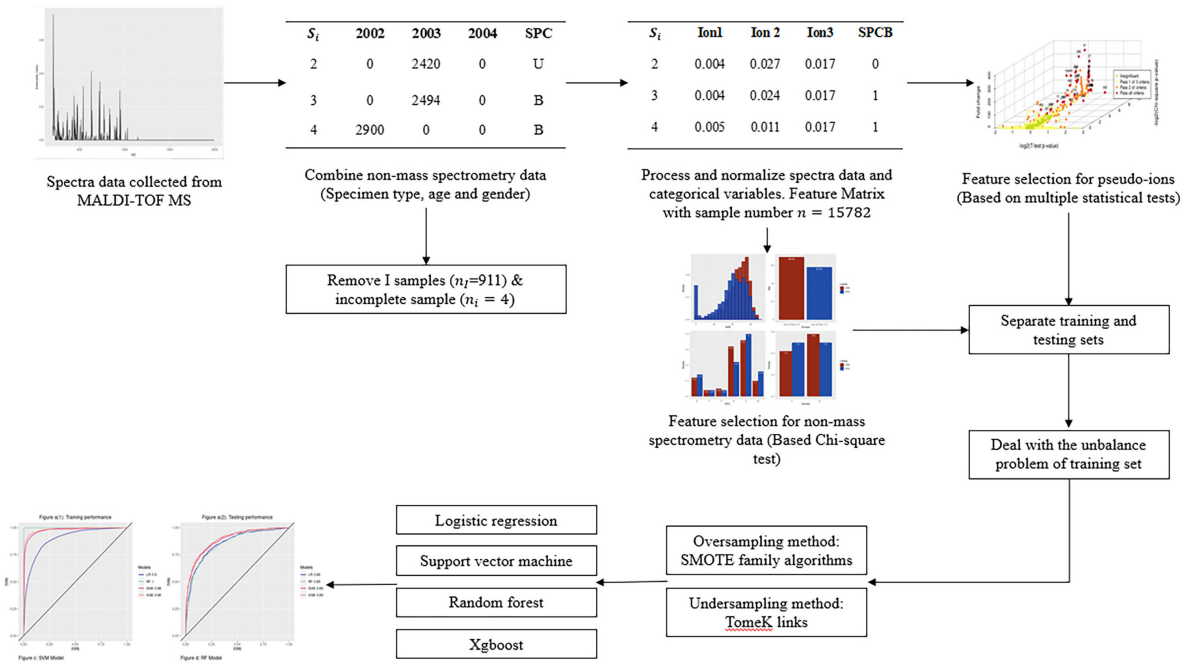


FIGURE 1 | Flow chart of the whole study, including sample collection, data cleaning and processing, feature selection, unbalance problem treatment, and model construction and comparison.

matrix was scaled as

$$\overrightarrow{P_k'} = \frac{\overrightarrow{P_k} - \overrightarrow{\overline{P_k}}}{\sigma\left(\overrightarrow{P_k}\right)}$$

where $\overrightarrow{\overline{P_k}}$ and $\sigma\left(\overrightarrow{P_k}\right)$ denote the mean and standard deviation of pseudo-ion k in order.

For non-spectrometric features, we grouped the specimen carriers (SPC) into six major types, which included blood sample (B), body fluid sample (F), wound sample (W), respiratory sample (R), urinary sample (U), and other types (O) (**Supplementary Table 1**). Continuous non-spectrometric feature age was transformed into an ordinal feature with six age groups for the convenience of statistical analysis (**Supplementary Table 2**). Labels of SPC after grouping and gender were directly used in statistical analysis. Besides, the modeling procedure included dummy variables generated from SPC, gender, and the scaled continuous feature age.

After data processing, the final feature matrix for modeling was formed by dummy variables of SPC, gender, and scaled age, and the pseudo-ion matrix.

Methods for Statistical Tests and Feature Selection

The Chi-square test was used to determine the impact of non-spectrometric characteristics on the CISKP and CIRKP groups. Non-spectrometric features with p-values less than 0.05 were considered as significant and would be included in the modeling.

The significance of pseudo-ions was mainly measured from three perspectives, and all pseudo-ions were tested with their intensity ratio before scaling: (1) the difference in mean values, (2) the difference in standard deviation, and (3) association to CIRKP. In the previous study, the t-test could be an option for determining the significance of the mean difference of log-transformed pseudo-ions (Wang et al., 2021). However, t-test pre-request normally distributed samples which was not the case in this study (**Supplementary Figures 1**, **2**). Therefore, the non-parametric Wilcoxon rank-sum test was used instead, coupling with the fold change selection on average intensity ratios to capture the information of mean shift of pseudo-ions with few observations. In this study, features with $\left|\log_2\left(fc\right)\right| \geq 1$ were considered as significant, where fc is the fold change value of the average intensity ratio between CIRKP and CISKP groups. As for the test of standard deviation difference, the traditional F-test for the equivalence of standard deviation was applied. Finally, the homogenous sample distributions between CIRKP and CISKP were tested by the Kolmogorov-Smirnov test (KS test) to directly measure the association between pseudo-ions and CIRKP. The significant level $\alpha = 0.05$ was used for a decision. Moreover, features were ranked by p-values (if the observation time of a feature was insufficient for testing, a p-value equal to 1 was set for that feature) of statistical tests, and features were also ranked by the fold change of average intensity ratio in the fold change selection.

The final ranks of pseudo-ions will overall consider all the test results above and take the average ranks of pseudo-ions as the final decision.

Clustering Method

The clustering method was based on the single-cell clustering approach, which was performed by R 4.1.1 with the assistance of the package "Seurat." We treat the intensity ratios over mass-to-charge ratio values as the expression level of genes.

Balancing Methods

For the modeling process in this study, data were first separated into training and testing sets by time, where the training set contained samples from June 1, 2013 to June 31, 2017 (85% of the total samples) and the remaining samples were used for performance testing. The training set only contained 3,613 CIRKP samples, which could cause an unbalancing problem.

Two balancing methods applied in this study were Tomek links (TKL) and Relocating Safe-Level SMOTE (RSLS). The TKL pair is defined as two samples from separate groups that are the closest neighbors (Tomek, 1976), which means there does not exist the third sample that the distances between the third sample and anyone of the TKL pairs are smaller than the distance between the TKL pairs. TKL can balance the training set by removing most class samples from TKL pairs. Besides, removing the whole TKL pairs can also alleviate the invasion problem.

RSLS is a modification of the Safe-Level SMOTE algorithm, which relocates the synthetic sample if the distance between the synthetic sample and a sample from the major class is less than the distance between that synthetic sample and its closest parent sample (Siriseriwan and Sinapiromsaran, 2016). This method considers the surroundings of synthetic samples and can provide a safer oversampling outcome for model training.

This study first removed both samples from all TKL pairs with Euclidean distance on the original feature space to relieve the invasion problem. After that, RSLS was applied to balance the CISKP and CIRKP classes. TKL was implemented with R function "TomeKClassif" in package "UBL," and RSLS balancing was implemented by calling function "RSLS" in package "smotefamily."

Machine Learning Models

Four popular machine learning models were tested in this study including logistic regression (LR), SVM, random forest (RF), and extreme gradient boosting (XGB). The general modeling processing in this study included (1) selecting significant features by multiple statistical methods and (2) constructing models with balancing methods and fivefold cross-validation (except the RF model). The final model included 480 features (472 pseudo-ions, 6 factors for SPC, 1 vector for age, and 1 factor for gender) selected by statistical methods. L1-regularization was applied to the LR model with $\mathbf{l}og\,(\lambda) = -\mathbf{4}.\mathbf{77}$ and finally 102 features were selected in the LR model. The testing performance of the SVM model was achieved by using radial kernel ("rbfdot" kernel in "kernlab" package) with penalty parameter $\mathbf{C} = \mathbf{1.5}$. There were 1,000 trees each with the random sampling size equal to 500 built for the RF model, but the RF model still suffered a serious overfitting problem. The XGB model was set to use "softmax" mode with the max depth equal to 3 and 527 iterations. All the models were implemented with R packages: "glmnet" for LR, "kernlab" for SVM, "randomForest" for RF, and "xgboost" for XGB. Predicting performances were measured by mean area under the receiver operating characteristic curve (AUC) and the accuracy rate of predicting CIRKP, and they are calculated with the assistance of the package "ROCR." Probability models allow flexible selection of probability cutoff. The optimal probability cutoffs (selected by balancing the specificity and sensitivity of models) of training and testing sets were both shown in this study. In addition, the predicting performance would be analyzed with the optimal cutoff of the training set since the optimal cutoff of the testing set was unobtainable in real antibiotic resistance prediction.

RESULTS

Insights for Specimen Information

An increase in the proportion of CIRKP was observed from 25.52% in 2014 to 29.07% in 2017 (**Figures 2A1,A4**). In total 7,556 (47.87%) and 8,226 (52.12%) samples were obtained from female and male patients. A total of 1,838 (41.51%) female samples and 2,590 (58.49%) male samples comprised the CIRKP group, whereas both female and male samples accounted for around 50% of the CISKP group, suggesting that men were probably more likely to be infected with CIRKP than women (**Figure 2A6**). Additionally, CIRKP was more likely to be diagnosed in people over the age of 60. The CIRKP group's average age was 11.15 years older than that of the CISKP group (**Figures 2A2,A3**). In the case of SPC, more than 60% of samples were collected from respiratory (R) or urinary (U) carriers. Additionally, respiratory (R) and other (O) samples accounted for a greater percentage of samples in the CIRKP group than in the CISKP group (**Figure 2A5**).

Spectrum data analyzed in this study were collected over the mass-to-charge ratio range from 2,000 to 20,000 m/z. To avoid the problems of magnitudes, the intensity ratio was used for analysis instead of the original intensity in this study. By comparing the average spectrum intensity ratio plot of the CISKP and CIRKP, CRSKP samples were found to have a lower intensity ratio over the lower region of the mass-to-charge ratio (2,000–3,000 m/z) and have a generally higher intensity ratio over the 3,000–7,000 m/z (**Figure 2B1**). However, no unique spectrum profile or spectrum peak could be observed directly from the average spectrum intensity plot. Moreover, the profiles of the average spectrum intensity ratio of both CISKP and CIRKP were shifting along with the time (**Figure 2C**). The intensity ratio of peaks at 2,069 m/z decreased from more than 4% to less than 2%. In contrast, peaks at 4,367, 5,382, and 6,291 m/z increased to more than 5% in the first 2 months of 2018. In addition, the numbers of spectrum peaks varied a lot among samples (**Figure 2B2**). Only 49 peaks were detected from the sample which is the minimum peak number. However, samples with over 900 peaks were also detected. Over 70% of samples contained peaks from 100 to 250. Both the spectrum profile shift and the wide range of peak numbers could indicate that specimens of different sub-strains of *K. pneumoniae* are collected, and their

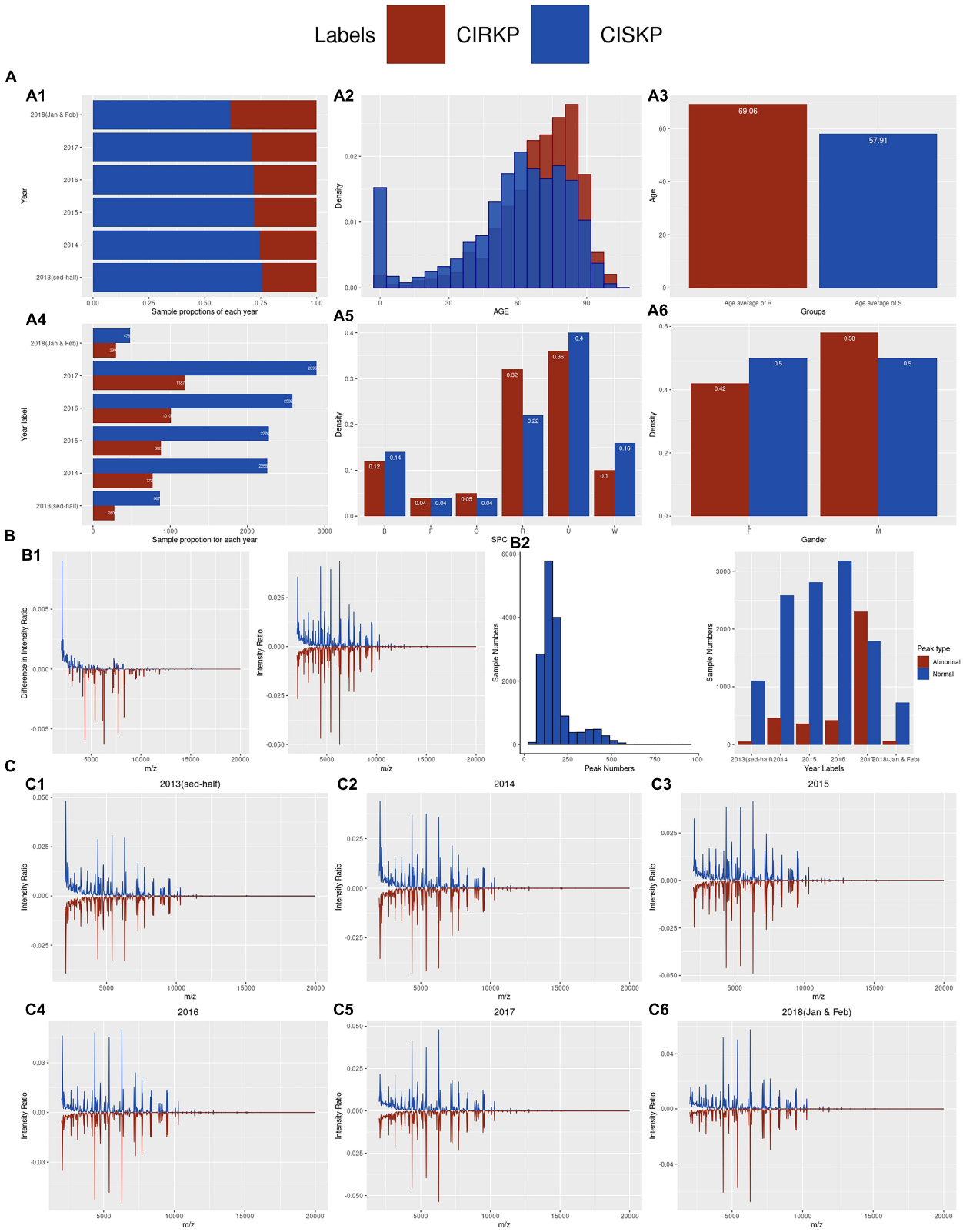


FIGURE 2 | Demographic of statistical information of the data. **(A1,A4)** Proportion of CIRKP samples in each year; **(A2,A3)** age information of samples; **(A5,A6)** number of samples of each SPC and gender in CIRKP and CISKP; **(B1)** overall average spectrum plot of CIRKP and CISKP; **(B2)** distribution of peak numbers. **(C1–C6)** average spectrum plots of CIRKP and CISKP by years.

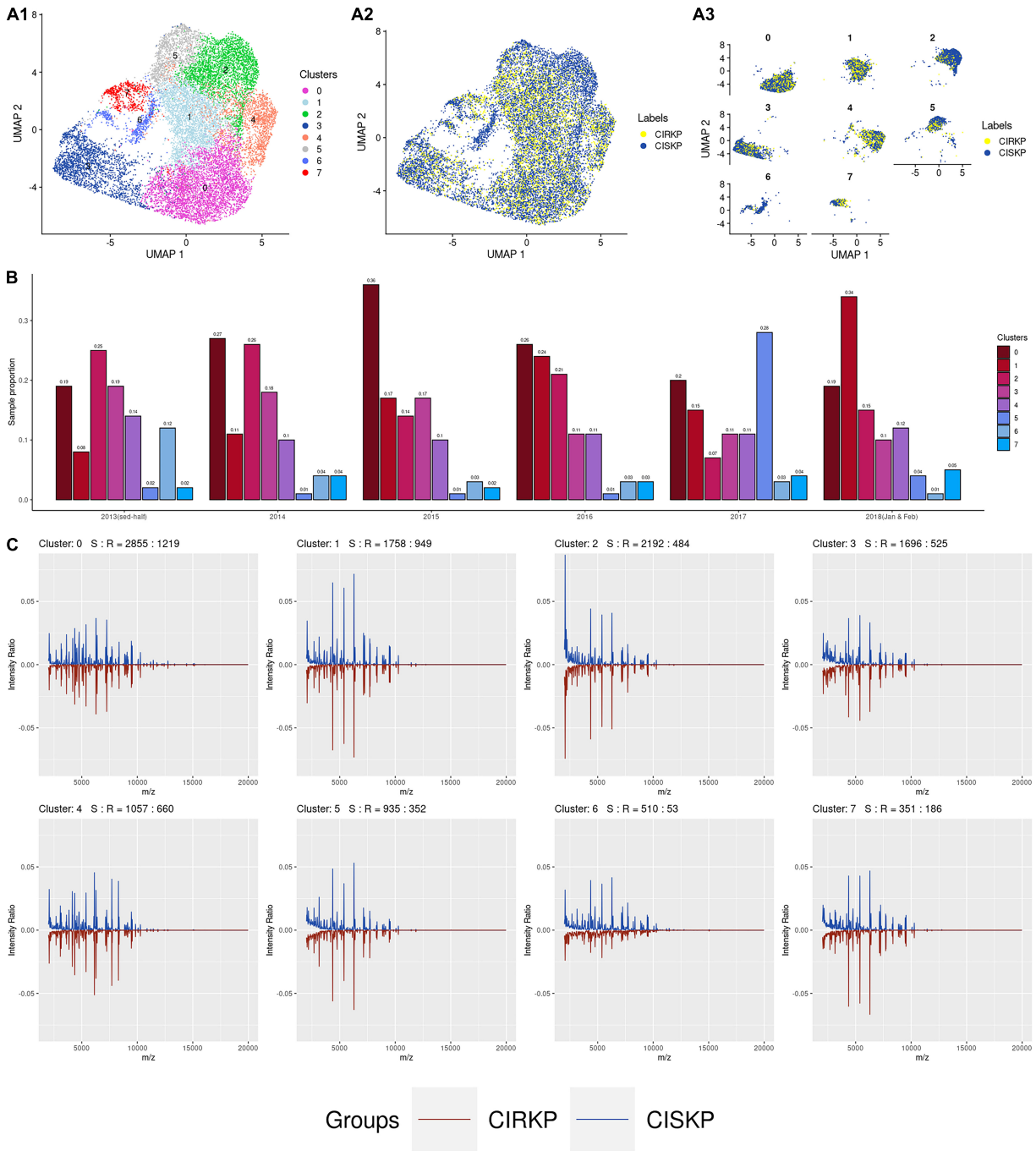


FIGURE 3 | Data visualization of clustering results. **(A1)** Distribution of 8 clusters; **(A2,A3)** distribution of CIRKP and CISKP samples in each cluster; **(B)** proportion of cluster in each year; **(C)** average spectrum plot of each cluster.

proportions were changing along with time. That was the main reason for the clustering analysis done in the following study.

Clustering Analysis

Due to the absence of standard MALDI-TOF MS spectrum plots for *K. pneumoniae* sub-strains, it was difficult to determine the true data composition. The clustering approach used in this study aimed at offering information about the composition of the samples under the assumption that bacteria from the same sub-strain have similar spectrum profile.

By setting the resolution parameter equal to 0.3, eight suspicious sub-strains were detected in this study. The cluster distributions and the distributions of CISKP and CIRKP samples under two-dimensional UMAP reduction are shown

TABLE 1 | Significance of non-spectrometry covariates.

		Sample population			
		CISKP	**CIRKP**	**Total number**	***P*-value**
Total numbers		9,201	3,424	12,625	
Gender (ratio %)					$< 2.2 \times 10^{-16}$
	Male	4,581 (49.7)	2,030 (59.3)	6,611	
	Female	4,620 (50.3)	1,394 (40.7)	6,014	
Age (ratio %)					$< 2.2 \times 10^{-16}$
	Infant	856 (9.3)	41 (1.2)	897	
	Children	50 (0.5)	12 (0.4)	62	
	Teenager	62 (0.7)	14 (0.4)	76	
	Youth	1,201 (13.1)	272 (7.9)	1,473	
	Middle-aged	3,715 (40.4)	1,167 (34.1)	4,882	
	Senium	3,317 (36.0)	1,918 (56.0)	5,235	
Specimen type (ratio %)					$< 2.2 \times 10^{-16}$
	B	1,330 (14.5)	400 (11.7)	1,730	
	F	351 (3.8)	133 (3.9)	484	
	W	1,490 (16.2)	373 (10.9)	1,863	
	R	2,081 (22.6)	1,148 (33.5)	3,229	
	U	3,610 (39.2)	1,276 (37.3)	4,886	
	O	339 (3.7)	94 (2.7)	433	

TABLE 2 | Top 15 significant pseudo-ions selected by statistical methods.

Rank	Pseudo-ion	m/z	Mean difference of peak intensity ratio (1^{0-4})	$\log_2$(fc)	Wilcoxon rank sum test	*F*-test	KS test	Observation times
1	**Pseudo-ion 136**	(4,700, 4,720)	10.22	0.88	1.24×10^{-38}	≤ 0.01	≤ 0.01	3,200
2	Pseudo-ion 5	(2,080, 2,100)	−50.75	−0.43	1.02×10^{-51}	≤ 0.01	≤ 0.01	12,528
3	Pseudo-ion 27	(2,520, 2,540)	−13.63	−0.41	1.73×10^{-64}	≤ 0.01	≤ 0.01	10,704
4	Pseudo-ion 95	(3,880, 3,900)	−3.51	−0.61	2.40×10^{-10}	≤ 0.01	≤ 0.01	3,358
5	Pseudo-ion 353	(9,040, 9,060)	−1.03	−0.77	5.60×10^{-05}	≤ 0.01	≤ 0.01	**478**
6	Pseudo-ion 1	(2,000, 2,020)	−13.52	−0.40	1.72×10^{-41}	≤ 0.01	≤ 0.01	10,727
7	Pseudo-ion 32	(2,620, 2,640)	−13.09	−0.40	1.01×10^{-34}	≤ 0.01	≤ 0.01	10,453
8	Pseudo-ion 293	(7,840, 7,860)	−0.26	−0.67	3.20×10^{-07}	≤ 0.01	≤ 0.01	**406**
9	Pseudo-ion 284	(7,660, 7,680)	−4.76	−0.72	1.87×10^{-03}	≤ 0.01	≤ 0.01	2,873
10	Pseudo-ion 11	(2,200, 2,220)	−13.45	−0.37	6.23×10^{-42}	≤ 0.01	≤ 0.01	10,405
11	Pseudo-ion 2	(2,020, 2,040)	−20.51	−0.38	9.96×10^{-34}	≤ 0.01	≤ 0.01	11,433
12	Pseudo-ion 4	(2,060, 2,080)	−97.53	−0.39	2.46×10^{-23}	≤ 0.01	≤ 0.01	12,318
13	**Pseudo-ion 367**	(9,320, 9,340)	0.01	1.11	3.86×10^{-04}	≤ 0.01	0.01	**285**
14	Pseudo-ion 50	(2,980, 3,000)	−12.41	−0.35	6.23×10^{-33}	≤ 0.01	0.01	10,574
15	Pseudo-ion 163	(3,240, 3,260)	−3.71	−0.84	2.46×10^{-09}	≤ 0.01	0.01	**768**

Mean difference is calculated by CIRKP-CISKP;fc represents the fold change value; fold change is calculated by CIRKP/CSIKP; total number of training samples: 13,414. Bold type values means the statistical quality of these pseudo-ions are relatively lower than other pseudo-ions since less samples are observed.

in **Figures 3A1,A2**. Intuitively speaking, clusters 2 and 6 are CISKP-dominant clusters with relatively lower CIRKP proportions compared to other clusters (**Figure 3A3**). Besides, CIRKP in clusters 2, 5, and 7 seems to be more separable from the CISKP sample than other clusters. But CIRKP and CISKP samples are highly mixed in the other clusters under two-dimensional UMAP reduction. However, no cluster shows a strong relation to the CIRKP or the CISKP group.

The trend of spectrum profile shifts is found from June 2013 to February 2018. It is worth mentioning that cluster 1 has grown from the fifth cluster in 2013 to the biggest cluster in 2018. Combined with the average spectrum plot of cluster 1, we can preliminarily conclude that the increase in the proportion of cluster 1 is the primary cause for the rise in the average intensity ratio of peaks at 4,367, 5,382, and 6,291 m/z (**Figures 3B,C**). At the same time, the proportion of cluster 0 first significantly increases from 2013 to 2015 and then gradually decreases to the same proportion level of 2013 in 2018. The proportion of cluster 2 keeps decreasing from about 25% of the whole year sample to 15%. As the result

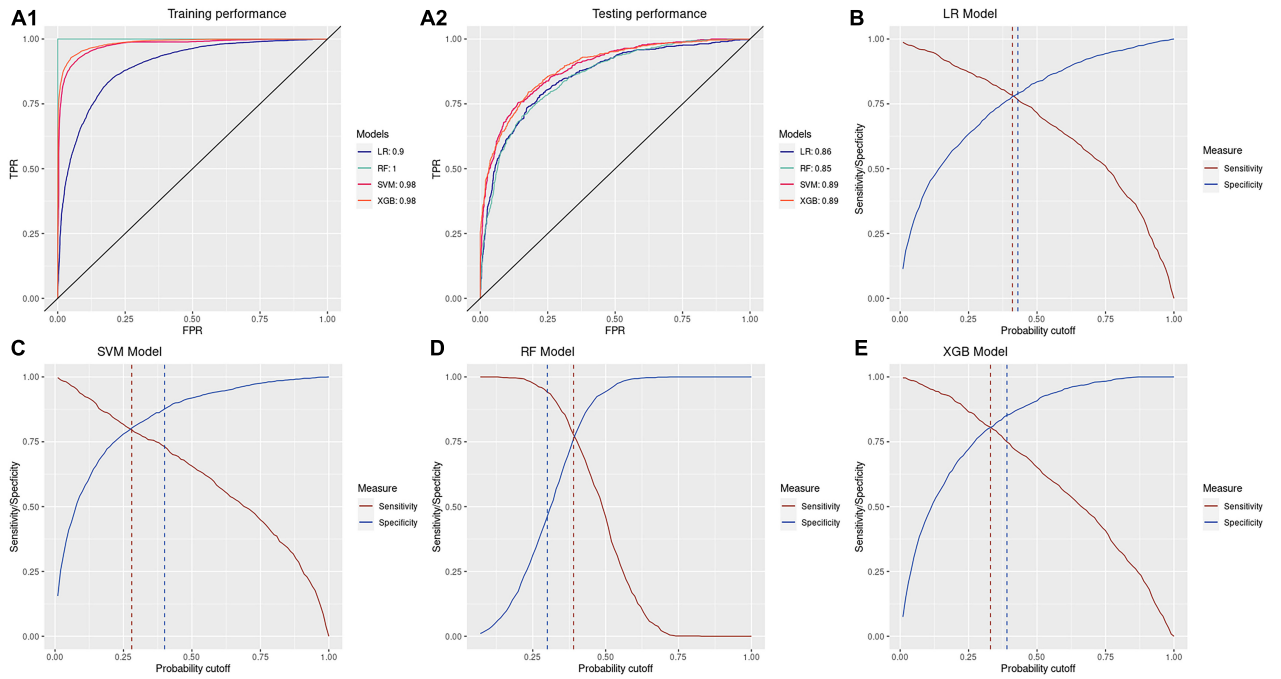


FIGURE 4 | Performance of models. **(A1,A2)** Training and testing ROC plots of four models; **(B–E)** optimal probability cutoff of training and testing set for four models.

TABLE 3 | Top 15 significant features selected by models.

	Features (observed sample number)														
	Rank 1	**Rank 2**	**Rank 3**	**Rank 4**	**Rank 5**	**Rank 6**	**Rank 7**	**Rank 8**	**Rank 9**	**Rank 10**	**Rank 11**	**Rank 12**	**Rank 13**	**Rank 14**	**Rank 15**
LR	AGE (13,414)	GEN (13,414)	U (5,218)	PI138 (9,756)	R (3,408)	PI172 (6,666)	PI307 (4,831)	PI228 (1,694)	PI158 (13,137)	PI495 (1,358)	PI103 (3,940)	PI203 (551)	PI290 (2,367)	PI287 (4,403)	PI92 (8,767)
SVM	PI154 (3,324)	AGE (13,414)	PI36 (3,200)	PI226 (5,434)	PI494 (2,858)	PI171 (9,543)	PI127 (1,739)	PI306 (5,431)	PI273 (606)	PI103 (3,940)	PI1495 (1,358)	PI230 (7,252)	PI367 (285)	PI198 (4,580)	PI52 (10,181)
RF	AGE (13,414)	PI171 (9,543)	R (3,408)	PI154 (3,324)	GEN (13,414)	PI136 (3,200)	PI208 (5,405)	PI91 (8,789)	PI288 (12,122)	PI165 (9,641)	PI316 (9,601)	PI226 (5,434)	PI108 (9,883)	PI286 (12,699)	PI306 (5,431)
XGB	R (3,945)	AGE (13,414)	GEN (13,414)	PI171 (9,543)	PI136 (3,200)	PI54 (3,324)	PI208 (5,405)	PI91 (8,789)	PI226 (5,434)	PI306 (5,431)	PI31 (9,619)	PI165 (9,641)	PI266 (5,375)	PI316 (9,601)	PI288 (12,112)

PI, pseudo-ion; GEN, gender; R, SPC-R; U, SPC-U.

of proportion changes of clusters 0, 1, and 2, spectrum peak at 2,069 m/z significantly decreases from 2013 to 2015, then increases a little bit in 2016, and sharply drops to less than 2% in 2018. Moreover, the proportion cluster 5 is found to be abnormally high only in 2017. The majority of cluster 5 are those with abnormally high number of spectrum peaks (peak number 250). Besides, most of the samples with abnormally low peak numbers (peak number 100) are found in cluster 1 (**Supplementary Figure 3**). For the other clusters, their proportions do not change significantly during the time of specimen collection.

After clustering analysis, it is normal to create classification models for each cluster. However, models trained by samples from all clusters were finally selected rather than models for each cluster. The reasons for this decision include:

1. The clusters in this study only represent suspicious sub-strains of *K. pneumoniae* without any additional support materials, implying that they are unreliable for modeling.
2. Not all cluster-based models outperform the overall model.
3. Most importantly, cluster-based models are unable to handle new samples from unknown clusters.

Feature Selection

For feature selection, statistical methods were applied on the training set to select associated non-spectrometric features and rank the significance of pseudo-ions.

The Chi-square test was used to determine the statistical significance of non-spectrometric variables using the original gender data as well as the age and SPC data after grouping (**Table 1**). The results of the tests indicated that CIRKP is

TABLE 4 | In cluster performance of the general model.

	Training (AUC %)				Testing (AUC %)			
Clusters	**LR**	**SVM**	**RF**	**XGB**	**LR**	**SVM**	**RF**	**XGB**
Overall	**89.69**	**97.61**	**100.00**	**98.50**	**85.58**	**88.86**	**85.44**	**89.08**
0	90.61	98.00	100.00	98.84	82.89	85.70	84.27	85.95
1	90.56	97.58	100.00	98.82	84.82	87.70	86.34	89.80
2	89.25	97.55	100.00	98.24	92.16	93.72	90.74	91.03
3	89.14	97.47	100.00	98.41	**80.05**	84.12	**79.41**	83.15
4	89.01	97.45	100.00	98.28	80.84	87.31	81.43	86.70
5	88.26	96.73	100.00	98.55	88.95	91.44	87.94	91.31
6	87.10	96.70	100.00	97.22	**50.00**	92.59	**74.07**	**57.41**
7	91.15	97.63	100.00	98.73	85.30	88.67	87.98	93.81

Bold type AUC value of cluster 0–7 shows poor performances of machine learning models on those clusters.

TABLE 5 | Confusion matrix of each model with optimal probability cutoff of training set.

Real predicted		**Cluster 0**		**Cluster 1**		**Cluster 2**		**Cluster 3**	
		CIRKP	**CISKP**	**CIRKP**	**CISKP**	**CIRKP**	**CISKP**	**CIRKP**	**CISKP**
LR	CIRKP	127	62	166	97	15	6	38	35
	CISKP	41	206	29	167	5	91	24	140
SVM	CIRKP	117	36	156	49	16	2	36	18
	CISKP	51	232	39	215	4	95	26	157
RF	CIRKP	**160**	**160**	**192**	**209**	**17**	**19**	**49**	**69**
	CISKP	**8**	**108**	**3**	**55**	**3**	**78**	**13**	**106**
XGB	CIRKP	105	37	161	57	14	5	35	22
	CISKP	63	231	34	207	6	92	27	153
		Cluster 4		**Cluster 5**		**Cluster 6**		**Cluster 7**	
		CIRKP	**CISKP**	**CIRKP**	**CISKP**	**CIRKP**	**CISKP**	**CIRKP**	**CISKP**
LR	CIRKP	91	46	159	57	**0**	**3**	28	20
	CISKP	22	103	61	413	**1**	**51**	8	55
SVM	CIRKP	85	27	156	37	**0**	**2**	30	21
	CISKP	28	122	64	433	**1**	**52**	6	54
RF	CIRKP	**111**	**129**	**202**	**187**	**0**	**11**	**36**	**49**
	CISKP	**2**	**20**	**18**	**283**	**1**	**43**	**0**	**26**
XGB	CIRKP	93	36	169	50	**0**	**2**	34	20
	CISKP	20	113	51	420	**1**	**52**	2	55
		LR		**SVM**		**RF**		**XGB**	
REF		**CIRKP**	**CISKP**	**CIRKP**	**CISKP**	**CIRKP**	**CISKP**	**CIRKP**	**CISKP**
	CIRKP	624	326	596	192	**767**	**825**	611	229
	CISKP	191	1,226	219	1,360	**48**	**727**	204	1,323

RF model is severely overfitted to CIRKP group. The in-cluster performance of cluster 6 is acceptable but low AUC value is caused by insufficient positive test samples.

associated with all non-spectrometric features. As a result, all dummy variables created from gender and SPC, and the actual age data, would be included in modeling.

After the construction of pseudo-ions, significance testing was performed using pseudo-ion vectors. Generally speaking, 18 pseudo-ions and 66 pseudo-ions are found uniquely in CIRKP and CISKP samples. However, only pseudo-ion 434 has been observed 12 times uniquely in CIRKP groups; the other unique pseudo-ions are all observed less than 5 times. Thus, the statistical power of those unique observed pseudo-ions is not guaranteed, and they will not be considered as highly significant features. Wilcoxon rank-sum test and average fold change selection were served for the significance check of mean values. The majority of the 209 pseudo-ions selected by the Wilcoxon rank-sum test concentrate at relatively low-intensity regions with high observation times, but the fold changes of those pseudo-ions are usually insignificant. In contrast, fold change selection selected pseudo-ions of highly different average intensity ratios, usually with lower observation times compared to pseudo-ions selected by the Wilcoxon rank-sum test. *F*-test and KS-test selected 303 and 208 pseudo-ions, respectively, and these pseudo-ions are highly overlapping with the pseudo-ions selected by the Wilcoxon rank-sum test. In conclusion, only pseudo-ions 367 and 407 passed all the selection criteria, and 154 pseudo-ions pass all tests except fold change selection.

The ranks of pseudo-ions were calculated by their average ranks of each selection criterion. The top 15 significant pseudo-ions selected by multiple statistical methods are shown in **Table 2**. Pseudo-ions that pass one of the four selection criteria were considered significant for modeling and would be used for model construction.

Classification Result

All four models showed high accuracy in detecting antibiotic resistance of patients. Both SVM and XGB have high AUC values of 0.89 on the testing set with specificities and sensitivities of 0.80 under the optimal probability cutoff of the testing set (**Figure 4**). Under the optimal cutoff of the training set, XGB and SVM could achieve accuracies of 0.82 [95% CI: (0.80, 0.83)] and 0.83 [95% CI: (0.81, 0.84)] of predicting CIRKP and CISKP samples in the testing set. The testing AUC of LR and RF is around 0.86 and 0.85, respectively. However, due to the severe overfitting problem of RF, the predicting performance under the optimal probability cutoff of the training set is highly unbalanced, which achieves an extremely high sensitivity of over 0.94 but low overall accuracy of 0.63 and unacceptable specificity of 0.47. Compared to RF, LR performs more stably. The gap between optimal cutoff of training and testing sets for LR is 0.02, which means LR is the only one of four models that could achieve both predicting sensitivity and specificity around 0.76 under the optimal cutoff of the training set time. However, since SVM and XGB could also achieve sensitivities of 0.73 and 0.75 and specificity of 0.88 and 0.85 simultaneously, they are considered slightly better model choices than LR.

Four models perform well on identifying CIRKP. We were interested in the differences between significant features selected by models and statistical methods (**Table 3**). The absolute value of coefficients directly ranked the feature significance of LR. The mean of decreasing in the Gini index was used to rank features that construct RF. The feature ranks for SVM and XGB models were created by constructing new models without target features

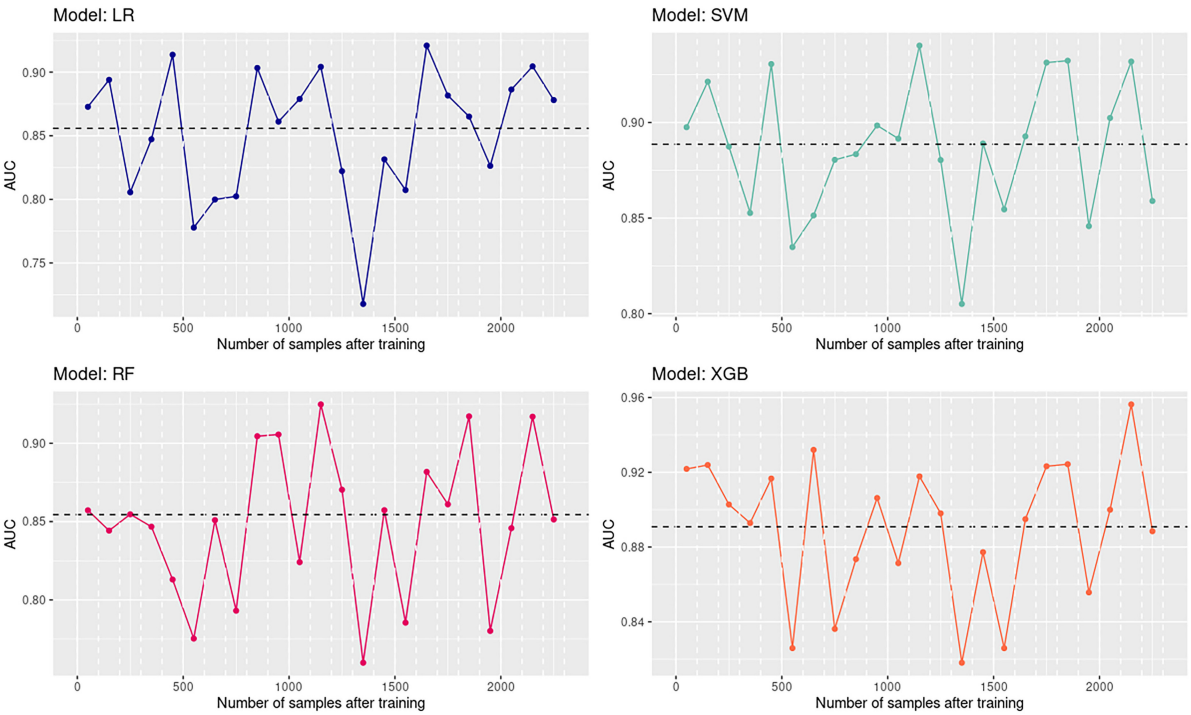


FIGURE 5 | Time influence on the performance of four models.

and calculating the decrease in testing AUC. For the results, only pseudo-ions 136 and 367 are considered the top 15 significant pseudo-ions by both models and statistical methods. Pseudo-ion 171 is also considered as one of the top 15 features by three models except for LR. Meanwhile, both LR and statistical method ranked pseudo-ion 171 at about 70th. Considering the 49 amino acids protein AcrZ subunit of AcrAB efflux pump whose expression change is proved to be one of the mechanisms of ciprofloxacin resistance, pseudo-ion 171 may be a representation of the expression level of AcrZ (Pakzad et al., 2013; UniProt Consortium et al., 2021). The statistical rank for features is very different from their real contribution in models. Statistical methods and each model select less than 10 identical top 100 features, respectively. However, feature contributions in RF and XGB models are highly consistent. Sixty-two identical features are found among their top 100 features. This observation could indicate that a better selection method or feature ranking system may help improve the classification outcome.

As for the non-spectrometric features, age is the most significant feature selected by models, constantly ranked as the top 2 significant features. Besides, gender is another highly significant feature ranked in the top 3 by LR, RF, and XGB models. These results show that patients' basic information may also contribute a lot to the real antibiotic resistance identification problems. Among all SPC types, only respiratory SPC (SPC-R) is considered one of the top 5 features of all models except SVM. Urinary SPC (SPC-U) is the third critical feature in LR, and it also ranked in the top 100 in the other models. Excluding SPC-U and SPC-R, other SPC types are not incredibly important for models. Besides, categorical variables are less popular for SVM compared to the other three models.

DISCUSSION

Analysis of In-Cluster Performance

Instead of constructing cluster-based models, four overall models constructed in this study utilized data from all clusters as training and testing samples. The overall performance of models showed high accuracy, and we wondered whether these models were competent for predicting antibiotic resistance for multiple suspicious sub-strains of *K. pneumoniae* (**Table 4**).

Four models can manage the in-cluster predicting task with acceptable performance most of the time. All four models exhibit high testing performance on predicting CIRKP in clusters 2, 5, and 7. This observation is consistent with the clustering analysis that these three clusters are more separable than other clusters. Furthermore, the outstanding performance on cluster 5 indicates that samples with abnormally high peak numbers are separable, and it is unnecessary to remove those samples from analysis. Compared to the other three models, SVM shows high stability in handling classification tasks on different clusters with the lowest AUC value of 0.84. Nevertheless, poor performances could also be found in predicting CIRKP on clusters 3 and 6. The predicting AUC value of LR and XGB are only 0.50 and 0.57 on cluster 6, but the training AUC achieves 0.87 and 0.97, respectively, which seems that both LR and XGB cannot perform the classification task better than a model performs random

selection. The predicting AUC of RF on cluster 6 also significantly drops 0.74. However, the low AUC values on cluster 6 are not due to the failure of models. By looking into the confusion matrixes under the optimal probability cutoff of the training set for each cluster (**Table 5**), only 1 CRIKP sample is found in the testing set of cluster 6. That means the AUC value of cluster 6 is unreliable. All models could achieve overall accuracy of over 0.92 except RF. As for model performances on cluster 3, LR, SVM, and XGB all show high specificity but low sensitivity. Even for RF, which achieves an overall sensitivity of 0.94, it cannot excellently classify CIRKP samples in cluster 3. The relatively inferior performance on cluster 3 provides evidence for the potential influence of sub-strains on the model.

In conclusion, four overall models constructed in this study can manage the classification task of CIRKP for most suspicious sub-strains, but the potential risk of sub-strain effect still exists. Determining the antibiotic resistance results by the optimal cutoff of each suspicious sub-strain instead of using the overall optimal cutoff of the training set may alleviate effect. However, a rigorous sub-strains identification process and solution to managing newly observed sub-strains should be set up.

Time Influence

Time influence on the predicting performance was considered by analyzing the fluctuation of the AUC value as the testing time period got far away from the last training samples. The first 2,300 samples in the testing sets were grouped into 23 groups by time, with 100 samples in each group. After that, the predicting AUC of each testing group is calculated and plotted in **Figure 5**. The expected decreasing trend in model performance along with time is not observed. Moreover, the same trend of the AUC changes is found among four models, which means the predicting performance of models is highly related to the separability of current samples. That emphasizes that even machine learning models with high classification accuracy in the past may fail in the current antibiotic resistance identification task. Reasons for the time effect could also include that the overall model cannot perform well on all the potential sub-strains, but the sample composition varies along with time.

Four models constructed in this study show high accuracy in identifying CIRKP. Especially SVM and XGB, which perform stably on all clusters and exhibit AUC greater than 0.8 for all the testing time periods. However, the stability against time effect could also be an essential criterion for clinically used antibiotic resistance predicting models.

IMPORTANCE AND CONCLUSION

The rising prevalence of CIRKP has increased the risk of incorrect selection of initial antibiotic treatment. The purpose of this study is to develop machine learning-based models for identifying antibiotic resistance of CIRKP using data from a longitudinal study done from June 2013 to February 2018. The use of large-scale data ensured the statistical quality of the selected biomarkers. Significant differences between the CISKP and CIRKP of a few pseudo-ions in the high mass-to-charge range were also detected, but their observation times were insufficient to draw a firm conclusion. Both statistical approaches and modeling algorithms recommended expanding the training set to conduct reliable statistical results for all the pseudo-ions. That indicated the need for a systematic and comprehensive MALTOF-MS database. Additionally, models were trained and evaluated in this work utilizing spectrum data from June 2013 to June 2017 and extra non-spectrometric information. Future samples were used to replicate the real-world antibiotic resistance prediction issue. Clustering analysis and the effect of time on model performance were implemented to provide potential information of suspicious sub-strains and demonstrate overall models' problems. While the model's performance does not meet the clinical standard, the findings of this investigation confirm the potential usefulness of the machine learning-based approach for antibiotic identification.

Limitations of this study include two main points, which are (1) the models may not be able to generalize to other bacterial species and antibiotics, and (2) the absence of biological validations. Protein analysis may be used to confirm the biological importance of selected features, and thorough real strain analysis, rather than the suspect sub-strain analysis used in this work, would give more helpful information. However, it is challenging to verify accurate sub-strains information in this study due to the data limitations. Additionally, other advanced machine learning techniques, such as deep learning and integrated prediction of multiple models, may enhance prediction accuracy.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

T-YL and ZW contributed to the study concept and design. H-YW and J-JL assisted in the data acquisition. CW and ZW carried out the data analysis. CW was response for drafted the manuscript. C-RC, J-TH, and J-JL helped conceive the study. All authors contributed to the data interpretation and revised the manuscript.

FUNDING

This work was supported by the Guangdong Province Basic and Applied Basic Research Fund (2021A1515012447), National Natural Science Foundation of China (32070659), the Science, Technology and Innovation Commission of Shenzhen Municipality (JCYJ20200109150003938), Ganghong Young Scholar Development Fund (2021E007), Shenzhen-Hong Kong Cooperation Zone for Technology and Innovation (HZQB-KCZYB-2020056), Chang Gung Memorial Hospital (CMRPG3L0401 and CMRPG3L0431), and the Ministry of Science and Technology, Taiwan (111-2320-B-182A-002-MY2).

ACKNOWLEDGMENTS

The authors sincerely appreciate the Warshel Institute for Computational Biology, the Chinese University of Hong Kong (Shenzhen) for financially supporting this research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.827451/full#supplementary-material

REFERENCES

Angeletti, S. (2016). Matrix assisted laser desorption time of flight mass spectrometry (MALDI-TOF MS) in clinical microbiology. *J. Microbiol. Methods* 138, 20–29. doi: 10.1016/j.mimet.2016.09.003

Arca-Suárez, J., Galán-Sánchez, F., Marin-Casanova, P., and Rodríguez-Iglesias, M. A. (2017). Direct identification of microorganisms from thioglycolate broth by MALDI-TOF MS. *PLoS One* 12:e0185229. doi: 10.1371/journal.pone.0185229

Ashurst, J. V., and Dawson, A. (2021). *Klebsiella pneumonia. StatPearls [Internet].* Available online at: https://www.ncbi.nlm.nih.gov/books/NBK519004/ (accessed November 23, 2021).

Burckhardt, I., and Zimmermann, S. (2018). Susceptibility testing of bacteria using MALDI-TOF mass spectrometry. *Front. Microbiol.* 9:1744. doi: 10.3389/fmicb.2018.01744

Clark, A. E., Kaleta, E. J., Arora, A., and Wolk, D. M. (2013). Matrix-assisted laser desorption ionization-time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. *Clin. Microbiol. Rev.* 26, 547–603. doi: 10.1128/CMR.00072-12

Clinical and Laboratory Standards Institute [CLSI] (2018). *Performance Standards for Antimicrobial Susceptibility Testing*, 28th Edn. Wayne: Clinical and Laboratory Standards Institute.

Florio, W., Tavanti, A., Barnini, S., Ghelardi, E., Lupetti, A., et al. (2018). Recent advances and ongoing challenges in the diagnosis of microbial infections by MALDI-TOF mass spectrometry. *Front. Microbiol.* 9:1097. doi: 10.3389/fmicb.2018.01097

Gaibani, P., Galea, A., Fagioni, M., Ambretti, S., Sambri, V., and Landini, M. P. (2016). Evaluation of matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of kpc-producing *Klebsiella pneumoniae*. *J. Clin. Microbiol.* 54, 2609–2613. doi: 10.1128/JCM.01242-16

Kumar, A., Ellis, P., Arabi, Y., Roberts, D., Light, B., Parrillo, J. E., et al. (2009). Initiation of inappropriate antimicrobial therapy results in a fivefold reduction of survival in human septic shock. *Chest* 136, 1237–1248. doi: 10.1378/chest.09-0087

Liu, X., Su, T., Hsu, Y. M. S., Yu, H., Yang, H. S., Jiang, L., et al. (2021). Rapid identification and discrimination of methicillin-resistant Staphylococcus aureus strains via matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 35:e8972. doi: 10.1002/rcm.8972

Martínez-Martínez, L., Pascual, A., and Jacoby, G. A. (1998). Quinolone resistance from a transferable plasmid. *Lancet* 351, 797–799. doi: 10.1016/s0140-6736(97)07322-4

Nordmann, P., and Poirel, L. (2005). Emergence of plasmid-mediated resistance to quinolones in Enterobacteriaceae. *J. Antimicrob. Chemother.* 56, 463–469.

Pakzad, I., Karin, M. Z., Taherikalani, M., Boustanshenas, M., and Lari, A. R. (2013). Contribution of AcrAB efflux pump to ciprofloxacin resistance in *Klebsiella pneumoniae* isolated from burn patients. *GMS Hyg. infect. Control* 8:Doc15. doi: 10.3205/dgkh000215

Patel, R. (2015). MALDI-TOF MS for the diagnosis of infectious diseases. *Clin. Chem.* 61, 100–111. doi: 10.1373/clinchem.2014.221770

Ruiz, J. (2003). Mechanisms of resistance to quinolones: target alterations, decreased accumulation and DNA gyrase protection. *J. Antimicrob. Chemother.* 51, 1109–1117. doi: 10.1093/jac/dkg222

Sanchez, G. V., Master, R. N., Clark, R. B., Fyyaz, M., Duvvuri, P., Ekta, G., et al. (2013). *Klebsiella pneumoniae* antimicrobial drug resistance, United States, 1998–2010. *Emerg. Infect. Dis.* 19, 133–6. doi: NODOI doi: 10.3201/eid1901.120310

Sandalakis, V., Goniotakis, I., Vranakis, I., Chochlakis, D., and Psaroulaki, A. (2017). Use of MALDI-Tof Mass Spectrometry in the battle against bacterial infectious diseases: recent achievements and future perspectives. *Expert Rev. Proteomic.* 14, 253–267. doi: 10.1080/14789450.2017.1282825

Singhal, N., Kumar, M., Kanaujia, P. K., and Virdi, J. S. (2015). MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front. Microbiol.* 6:791. doi: 10.3389/fmicb.2015.00791

Siriseriwan, W., and Sinapiromsaran, K. (2016). The effective redistribution for imbalance dataset: relocating safe-level SMOTE with minority outcast handling. *Chiang Mai J. Sci.* 43, 234–246.

Sogawa, K., Watanabe, M., Ishige, T., Segawa, S., Miyabe, A., Murata, S., et al. (2017). Rapid discrimination between methicillin-sensitive and methicillinresistant *Staphylococcus aureus* using MALDI-TOF mass spectrometry. *Biocontrol Sci.* 22, 163–169. doi: 10.4265/bio.22.163

Tang, W., Ranganathan, N., Shahrezaei, V., and Larrouy-Maumus, G. (2019). MALDI-TOF mass spectrometry on intact bacteria combined with a refined analysis framework allows accurate classification of MSSA and MRSA. *PLoS One* 14:e0218951. doi: 10.1371/journal.pone.0218951

Tomek, I. (1976). Two modifications of CNN. *IEEE Trans. Syst. Man Cybern.* 6, 769–772. doi: 10.1109/tsmc.1976.4309452

Tré-Hardy, M., Lambert, B., Despas, N., Bressant, F., Laurenzano, C., Rodriguez-Villalobos, H., et al. (2017). MALDI-TOF MS identification and antimicrobial susceptibility testing directly from positive enrichment broth. *J. Microbiol. Methods* 141, 32–34. doi: 10.1016/j.mimet.2017.07.012

Tumbarello, M., Viale, P., Viscoli, C., Trecarichi, E. M., Tumietto, F., Marchese, A., et al. (2012). Predictors of mortality in bloodstream infections caused by *Klebsiella pneumoniae* carbapenemase-producing *K. pneumoniae*: importance of combination therapy. *Clin. Infect. Dis.* 55, 943–950. doi: 10.1093/cid/cis588

UniProt Consortium, European Bioinformatics Institute, Protein Information Resource, and Swiss Institute of Bioinformatics [SIB] (2021). *Multidrug Efflux Pump Accessory Protein Acrz*. Available online at: https://www.uniprot.org/uniprot/W1HX74 (Accessed February 10, 2021)

van Belkum, A., Welker, M., Pincus, D., Charrier, J. P., and Girard, V. (2017). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry in clinical microbiology: what are the current issues ?. *Ann. Lab. Med.* 37, 475–483. doi: 10.3343/alm.2017.37.6.475

Wang, Z., Wang, H. Y., Chung, C. R., Horng, J. T., Lu, J. J., and Lee, T. Y. (2021). Large-scale mass spectrometry data combined with demographics analysis rapidly predicts methicillin resistance in Staphylococcus aureus. *Brief. Bioinform.* 22:bbaa293. doi: 10.1093/bib/bbaa293

Weis, C. V., Jutzeler, C. R., and Borgwardt, K. (2020). Machine learning for microbial identification and antimicrobial susceptibility testing on MALDI-TOF mass spectra: a systematic review. *Clin. Microbiol. Infect.* 26, 1310–1317. doi: 10.1016/j.cmi.2020.03.014

Zhou, X. Y., Ye, X. G., He, L. T., Zhang, S. R., Wang, R. L., Zhou, J., et al. (2016). In vitro characterization and inhibition of the interaction between ciprofloxacin

and berberine against multidrug-resistant *Klebsiella pneumoniae*. *J. Antibiot.* 69, 741–746. doi: 10.1038/ja.2016.15

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 12 April 2022
doi: 10.3389/fmicb.2022.853775

Rapid Antibiotic Resistance Serial Prediction in *Staphylococcus aureus* Based on Large-Scale MALDI-TOF Data by Applying XGBoost in Multi-Label Learning

Jiahong Zhang[1,2], Zhuo Wang[1,3], Hsin-Yao Wang[4,5], Chia-Ru Chung[6], Jorng-Tzong Horng[6,7], Jang-Jih Lu[4,8,9] and Tzong-Yi Lee[1,2]**

[1] Warshel Institute for Computational Biology, The Chinese University of Hong Kong, Shenzhen, China, [2] School of Life and Health Sciences, The Chinese University of Hong Kong, Shenzhen, China, [3] School of Life Sciences, University of Science and Technology of China, Hefei, China, [4] Department of Laboratory Medicine, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan, [5] Ph.D. Program in Biomedical Engineering, Chang Gung University, Taoyuan, Taiwan, [6] Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan, [7] Department of Bioinformatics and Medical Engineering, Asia University, Taichung, Taiwan, [8] Department of Medical Biotechnology and Laboratory Science, Chang Gung University, Taoyuan, Taiwan, [9] Department of Medicine, College of Medicine, Chang Gung University, Taoyuan, Taiwan

OPEN ACCESS

Edited by:
Antonella Lupetti, University of Pisa, Italy

Reviewed by:
Marina Oviaño, A Coruña University Hospital Complex (CHUAC), Spain
Liang Qiao, Fudan University, China

***Correspondence:**
Jang-Jih Lu
janglu45@gmail.com
Tzong-Yi Lee
leetzongyi@cuhk.edu.cn

Specialty section:
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

Received: *13 January 2022*
Accepted: *11 March 2022*
Published: *12 April 2022*

Citation:
Zhang J, Wang Z, Wang H-Y, Chung C-R, Horng J-T, Lu J-J and Lee T-Y (2022) Rapid Antibiotic Resistance Serial Prediction in Staphylococcus aureus Based on Large-Scale MALDI-TOF Data by Applying XGBoost in Multi-Label Learning. Front. Microbiol. 13:853775. doi: 10.3389/fmicb.2022.853775

Multidrug resistance has become a phenotype that commonly exists among *Staphylococcus aureus* and is a serious concern for infection treatment. Nowadays, to detect the antibiotic susceptibility, antibiotic testing is generated based on the level of genomic for cure decision consuming huge of time and labor, while matrix-assisted laser desorption-ionization (MALDI) time-of-flight mass spectrometry (TOF/MS) shows its possibility in high-speed and effective detection on the level of proteomic. In this study, on the basis of MALDI-TOF spectra data of discovery cohort with 26,852 samples and replication cohort with 4,963 samples from Taiwan area and their corresponding susceptibilities to oxacillin and clindamycin, a multi-label prediction model against double resistance using Lowest Power set ensemble with XGBoost is constructed for rapid susceptibility prediction. With the output of serial susceptibility prediction, the model performance can realize 77% of accuracy for the serial prediction, the area under the receiver characteristic curve of 0.93 for oxacillin susceptibility prediction, and the area under the receiver characteristic curve of 0.89 for clindamycin susceptibility prediction. The generated multi-label prediction model provides serial antibiotic resistance, such as the susceptibilities of oxacillin and clindamycin in this study, for *S. aureus*-infected patients based on MALDI-TOF, which will provide guidance in antibiotic usage during the treatment taking the advantage of speed and efficiency.

Keywords: MALDI-TOF MS, oxacillin resistance, clindamycin resistance, XGBoost, multi-label learning

INTRODUCTION

The multidrug resistance phenotype that occurred within *Staphylococcus aureus* is considered as one of the most intractable pathogenic features in the history of antibiotic chemotherapy (Hiramatsu et al., 2014). This feature refers to *Staphylococcus aureus*, which shows resistance to a set of antibiotics. oxacillin-resistant *S. aureus* (ORSA) has been increasing in importance as a leading cause of both nosocomial and community-acquired infections (Bell and Turnidge, 2002). Similar to penicillin and methicillin, oxacillin belongs to β-lactam drugs. The initial discovery on the mechanism of β-lactam drugs is the existence of penicillin-binding-proteins (PBPs), which are transpeptidases responsible for partial peptidoglycan construction on cell walls. The binding between penicillin and PBPs blocks the function of PBPs and creates the entry for penicillin. Gene blaZ was induced in bacteria encoding a β-lactamase enzyme, which opens up the β-lactam ring at the core of penicillin, preventing the binding to PBPs. Oxacillin resistance results from a new PBP, decreasing the affinity for oxacillin, though the β-lactam ring within the drug has been modified and stabilized. Clindamycin-resistant *Staphylococcus aureus* (CRSA) is free of the suppression in the virulence factors expression, which is originally regulated by clindamycin (Hodille et al., 2018). The mechanism of clindamycin is binding to the ribosome and inhibiting protein synthesis (Kehrenberg et al., 2005). Correspondingly, clindamycin resistance results from conformation change of ribosome induced by enzymes, which leads to the affinity decreasing (Reygaert, 2013).

Nowadays, to test antibiotic susceptibility, the workflow takes 24–72 h including disk diffusion. Basically, in the case of the low-efficiency treatment, patients infected by ORSA or CRSA are asked to do the test and wait for the detection result (Swenson et al., 2001), which causes a delay for the concise and precise treatment individually ranging from 24 to 72 h, though broad-spectrum empirical treatment would be conducted. Besides, long-time testing is not suitable for urgent patients and leaves the time lag for the probability of mutation.

In recent years, a huge number of arrays or kits emerged and be applied in clinical detection such as Velogene and MRSA-Screen, improving the detection time within 4 h (Louie et al., 2000). For instance, Velogene uses a chimeric probe aiming at the mecA gene within 90 min. Nevertheless, the high cost of detection kits and limited labor capacity, privacy policy restricts the application of genome detection. Proteomic of the resistant *S. aureus* is also a focus of identification based on the ion types and expression intensity generated by the spectra. Current antibiotic susceptibility tests have shortened the detection time within several hours besides *S. aureus* isolation and culture. Nevertheless, the time lag still exists the chance for resistance induction, which is calling for rapid detection and proteomic-based tests with statistics and computational algorithms. Specifically, with the availability of matrix-assisted laser desorption-ionization (MALDI)-time-of-flight mass spectrometry (TOF/MS), its fast generation speed and accurate fragmentation detection are the advantages as well as cross-species processing, which are longed for a long time to solve resistance detection.

Matrix-assisted laser desorption-ionization-time-of-flight mass spectrometry is a special kind of mass spectrometry technique that requires protein samples crystallized within the matrix for further ionization and detection, which can be applied to grasp the resistant characters besides antibiotic susceptibility testing (Lay, 2001; Croxatto et al., 2012). Each run of detection through MALDI-TOF only causes low cost within a few dollars within 5 min. Scientists have tried to combine the statistical analysis, computational method, even machine learning with the spectra information such as mass-to-charge (m/z) ratio and peptide intensity from the MALDI-TOF to differentiate sensitive and resistant *S. aureus* for several types of antibiotics (Wang et al., 2020). Through the combination of MALDI-TOF and machine learning, the classification model could be a guide to provide insight information into drug susceptibility during the clinical treatment and even show the potential of saving the antibiotic test in the ideal case.

The crucial consideration from both patients and doctors is that the computational model on the basis of the cohort representation and assumption lacks quality guarantee for individuals, which can be solved and ensured largely in the antibiotics susceptibility test. Specifically, the consideration is getting mitigated with a novel resistance information database called DRIAMS with huge-scale data, which collects at least 300,000 mass spectra with more than 750,000 antimicrobial resistances (Weis et al., 2022). Another limitation is that each existing classification model only refers to a specific type of antibiotic, which is not suitable and applicable for the multidrug-resistant *S. aureus* with the widespread multidrug-resistant phenotype, referring to being resistant to at least three classes of antibiotic mechanisms or three antibiotics based on the gene level (Schwarz et al., 2010). Thereby, to relieve the dilemma in cohort representation and size, this study recruited 26,852 patients infected by *S. aureus* in the Chang Gung Memorial Hospitals (CGMH) at the Linkou branch from 2013 to 2019. Antibiotic susceptibility tests on oxacillin and clindamycin had been conducted for the samples, and their *S. aureus* susceptibility status was mapped with their MALDI-TOF results as the labels. Besides, for the reproducibility, from 2015 to 2017, this study also recruited 4,963 patients as the validation test for the constructed model at the Kaohsiung branch. In our dataset, information, such as specimen type, sex, age, m/z, and peak intensity, is included for each sample. Two drug susceptibilities are combined in the form of a tuple as the label data. We aimed to construct a prediction model using the high-dimension and large-scale MALDI-TOF data to indicate the resistance for oxacillin and clindamycin in patients, which is not the typical case for multiresistance but breaking model mode for single susceptibility prediction. Meantime, this study applies the XGBoost algorithm in multi-label learning for fast and accurate serial resistance prediction. Over the long haul, the model could improve resistance detection, provide medication guidance, and be extended to serial antibiotic susceptibility tests.

MATERIALS AND METHODS

Overview of the Study

This study consists of sample data generation and prediction model construction (as shown in **Figure 1**). During the first step, the clinical specimen from recruited patients is cultured, and a single pathogen colony after isolation and incubation was treated with the MALDI-TOF MS spectra. During the second step, the mass spectra data generated by the MALDI-TOF from different samples went through the preprocessing and were used for the modeling of serial-drug resistance prediction. **Figure 1** presents the overlook of this study including the basic flow of the MALDI-TOF MS spectra and the process of model construction. Study details are introduced in the following sections.

Experiment Cohorts' Information, MALDI-TOF Preparation, and Processing

Two cohorts, Linkou and Kaohsiung, are used as the discovery and replication population, respectively, which is independent of each cohort. The Linkou cohort focused on the oxacillin and clindamycin resistance of *S. aureus* lasting from 2013 to 2019. We collected wound (W) swab specimens, respiratory tract (RT), sterile body fluid (SBF), blood (B), and urinary tract (UT) from patients from different departments during the data tracking. For those samples that showed resistance in the AST, CGMH cultured the clinical specimens, isolated bacterial pathogens from the samples, and did the antibiotic resistance profiling. **Table 1** presents the label information of the Linkou cohort, which is prepared for the discovery part (**Table 2**). Notably, 26,852 samples in total are marked with two labels after the antibiotic susceptibility testing to the oxacillin and clindamycin. One of the labels can be categorized as ORSA or oxacillin-sensitive *Staphylococcus aureus* (OSSA). The other one can be presented as CRSA or clindamycin-sensitive *Staphylococcus aureus* (CSSA). Meanwhile, more information, such as age and sex, was collected from each participant. In the Kaohsiung cohorts (**Table 2**), 4,963 samples were collected from 2015 to 2017 as another independent cohort and treated with the consistent processing procedures as the Linkou cohort.

TABLE 1 | Susceptibility information for clindamycin and oxacillin in the Linkou cohort.

Clindamycin \ Oxacillin	Susceptible	Resistant	Total
Susceptible	11,453	3,761	15,214
Resistant	1,539	10,099	11,638
Total	12,992	13,860	26,852

TABLE 2 | Susceptibility information for clindamycin and oxacillin in the Kaohsiung cohort.

Clindamycin \ Oxacillin	Susceptible	Resistant	Total
Susceptible	2,303	800	3,103
Resistant	288	1,572	1,860
Total	2,591	2,372	4,963

Besides the label information, more basic information for samples including specimen types, gender, and age is shown in **Figure 2**. In each subgraph, the sample composition under each category is presented. For **Figures 2A,B**, the subgraphs in both mainly share the same composition situation corresponding to the same horizontal coordinate (specimen types, gender, and age).

Specimens are treated individually and separately with corresponding methods for sample culture. Notably, 1.2 ml of 0.9% saline solution is added to rinse the W swab specimens. Following, transfer equivalently onto four kinds of culture media including blood plate agar, eosin methylene blue agar, Columbia nalidixic acid, and chocolate agar. As for blood, we used a blood culture kit (BD BACTEC^{TM} FX), which is for commercial use and from Becton, Dickinson and Company, to isolate pathogens. Following the positive blood culture bottle, we inoculated it on blood plate agar to regain single colonies. Sharing a similar

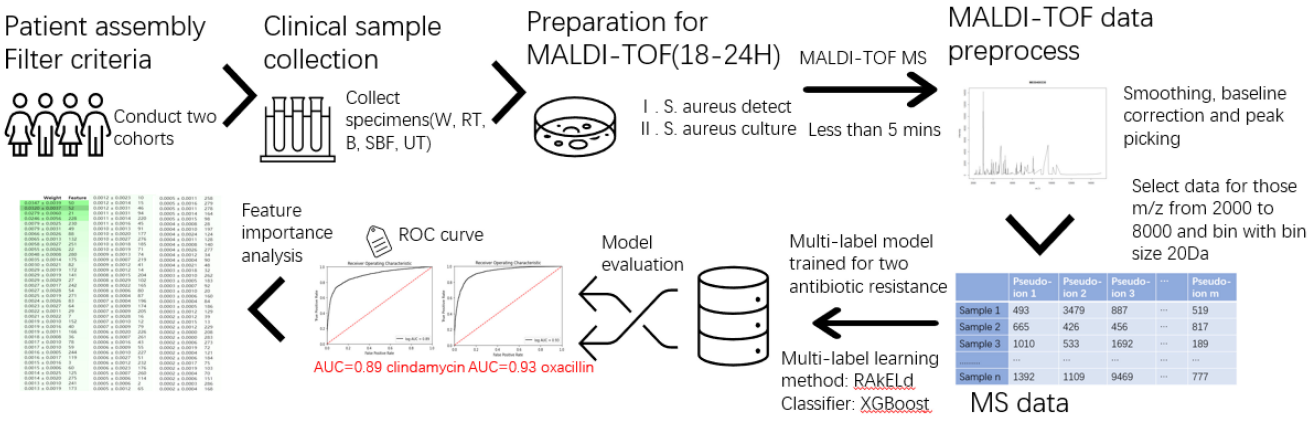


FIGURE 1 | Overview of the study. The flowchart mainly contains sample collection, matrix-assisted laser desorption-ionization-time-of-flight mass spectrometry (MALDI-TOF) process, and multi-label model construction and evaluation.

FIGURE 2 | Data composition information in isolated *Staphylococcus aureus* for Linkou cohort in subgraph **(A)** (discovery population) and Kaohsiung cohort in **(B)** (replication population). The turtle in the graph stands for clindamycin susceptibility and oxacillin susceptibility. R is for resistant, and S is for susceptible. **(A)** Specimen types, gender, and age map to multi-labels in the discovery population. **(B)** Specimen types, gender, and age map to multi-labels in replication population.

protocol, such as W swab specimens, sterile body fluid is added onto the four agars, same as the W swab specimens, and rinsed by liquid thioglycolate for microorganisms' isolation. After getting the culture prepared, agars and media were put into a 37°C CO_2 incubator for 18–24 h. After the culture, we selected single colonies on the agar plate for the analysis of MALDI-TOF mass spectrometry. The isolates were collected consecutively. One isolate was generally for one patient. If there were multiple isolates of the same species, the first isolate was used. With the identification of the *S. aureus* from the colonies, oxacillin and clindamycin susceptibility tests are applied to label the two susceptibilities to the colonies. The technique and reagents are originated from the cefoxitin disk (Clinical and Laboratory Standards Institute guideline)[1] for non-sterile specimens. For the sterile specimens including B specimens, the broth microdilution method is used as the resistance test.

Our cohorts were analyzed under MALDI-TOF MS (Microflex LT MALDI-TOF System, Bruker Daltonik GmbH). The operation requirement and processes were run under the manufacturer's guidance. Each step is as follows. (1) prepare a MALDI steel target plate, smear the colonies after culture with a thin film adding formic acid (1 μl, 70%), and get dried at 25°C, (2) prepare the matrix solution based on the guidance and kit (1% α-cyano-4-hydroxycinnamic acid in 50% acetonitrile containing 2.5% trifluoroacetic acid), (3) add the matrix solution to the film and get dried under room temperature, and (4) microflex LT MALDI-TOF analyzer was operated to analyze the samples (linear ionization mode; accelerating voltage, 20 kV; nitrogen laser frequency: 60 Hz; 240 laser shots). In the end, we generated the raw MALDI-TOF data, whose m/z ratio ranged from 2,000 to 20,000 Da.

MALDI-TOF Data Preprocessing and Pseudo-Ion Peak Intensity Matrix Generation

In the part of raw data preprocessing, three techniques were used to treat the data by order. An external calibration (Bruker Daltonics Bacterial Test Standard) was applied as the first step. Later, peak smoothing was performed using the Savitzky–Golay filter, and baseline correction was performed using the Top-hat filter. Peaks with a signal-to-noise ratio were set larger or equal to 2 for further analysis.

The preprocessed data based on the raw data consist of two categories of the variable for each sample: m/z and peak intensity. For the further preprocess, first, filter the unqualified mass spectra that the number of peaks is lower than 100 or larger than 200. Subsequently, by considering the sparsity of MS data when the m/z ratio is larger than 8,000 Da and signal regarding phenol-soluble modulin (PSM)-mec was studied earlier, which is a peptide with 2,415 m/z encoded by resistance gene, mecA (Josten et al., 2014), the MS data are extracted for each sample based on the range of m/z from 2,000 to 8,000 Da. Meanwhile, to minimize the impact of peak shift caused by different fragmentation results due to the initial point, the window size of 20 Da is considered to modify the data and transfer the m/z ratio into pseudo-ions. Specifically, the first pseudo-ion includes the intensity for the m/z ratio ranges from 2,000 to 2,010 Da as same as the last pseudo-ion. Other pseudo-ions between them stand for an interval lasting for 20 Da. In the end, a total of 301 pseudo-ions are generated, and the intensity corresponding to a pseudo-ion is the intensity sum within the interval. The intensity of the ith ($i = 1,\ldots,301$)

[1]www.clsi.org

pseudo-ion for one sample can be calculated as follows:

$$intensity'_{(i)} = \sum_{j=1}^{interval\,(i)} intensity_{(j)}$$

where in one sample, $intensity'_{(i)}$ stands for the intensity corresponding to ith pseudo-ion. Interval (i) refers to the m/z ratios within the interval, and $intensity_{(j)}$ is the intensity for a specific m/z ratio.

Multi-Label Classification

Multi-label learning studies the problem where each example is represented by a single instance while associated with a set of labels simultaneously (Zhang and Zhou, 2014). In this study, the pseudo-ion-intensity data are the observation data, and the results of the susceptibility test for oxacillin and clindamycin are assigned as the label data. All the multi-label learning algorithms are from scikit-multilearn 0.2.0 (Szymański and Kajdanowicz, 2017), a library for multi-label classification built on top of the scikit-learn ecosystem, using Python 3.68.

Binary Relevance

Binary relevance is the most intuitive idea to deal with multi-label prediction. It treats the multi-label separately by considering multiple independent binary classifications for each label instead of viewing it as a group of labels. Like in this study, for the binary relevance, it needs to train two models, and the output is the union of two separated predictions.

Classifier Chain

Classifier chain is the improved transformer of binary relevance by the construction of a Bayesian conditioned chain. Similar to the binary relevance, the classifier chain treats each label as a separated classier but not independent. Although the first classier is only trained using the input data (observation), the classifiers after are trained on the input space and all previous classifiers in the chain based on the Bayesian chain rule by order.

Lowest Power Set

Unlike the previous two methods, the lowest power set is to transform a multi-label problem into a multi-class problem. Like this study, for 2 labels totally, it will eventually transform into a 4-class classifier.

Quality Measures for Multi-Label Model

Two measurements are applied to evaluate the multi-label model in this study. The first measures are the hamming loss, which stands for the proportion of the incorrect prediction for all labels among the whole samples. The other measurement is the accuracy score. It means the fraction of samples for those prediction sets that exactly match the real label sets.

Logistic Regression

Inside the multi-label algorithm, the classifier needs to be defined. Logistic regression is a common classifier to predict the resistance in the biological field (Moradigaravand et al., 2018). In this study, the discovery samples, the Linkou cohort, are used in the model training, while the Kaohsiung cohort is responsible for the independent test. The logistic regression (LR) model is realized using the Python package sklearn. Grid search is applied for the parameter tuning based on the criteria of the area under the receiver operating characteristic (ROC) curve by the adjustment of parameters including the penalty, C-value, and solver. Each model during the tuning is evaluated by the 5-fold cross-validation. For the tuned model, using L1 normalization as the penalty, 1 for the C-value, and liblinear, a library for large linear classification, as the solver is the tuned parameters. The area under the curve (AUC) will be applied to evaluate the training model in the replication cohort. The model training and parameter tuning are achieved in the Python package, scikit learn. The presentation of the ROC curve is generated from the Python package, Matplotlib.

XGBoost

XGBoost is a scalable machine learning system for the tree boost, offering parallel tree boosting (Chen and Guestrin, 2016). Choosing XGBoost as the classifier in the multi-label model, such as the LR above, the Linkou cohort is treated as the training data, and the Kaohsiung cohort is used for the independent test by orders. Package xgboost from Python is applied to realized XGBoost model. Parameters shown in **Table 3** were tuned through grid search. The result evaluation of the model is the same procedure as the LR.

Permutation Importance

Permutation importance is a technique used to generate the feature importance for the trained model. It is defined as the decrease of significance P-values for each feature when the value is randomly shuffled (Altmann et al., 2010).

TABLE 3 | Parameters tuned for XGBoost under multi-label learning.

Parameter	Function	Tuned result
max_depth	Maximum depth of a tree	3
min_child_weight	Minimum sum of weight for a child	1
Gamma	Minimum loss requirement for node partition	0
subsample	Subsample ratio within the training samples	0.6
colsample_bytree	Subsample ratio of columns when constructing each tree	0.6

TABLE 4 | Model evaluation in multi-label ensembles using LR and XGBoost correspondingly.

Criteria \ Ensembles	BinaryRelevance LR (XGBoost)	ClassifierChain LR (XGBoost)	Lowest Power set LR (XGBoost)
Hamming loss	0.2023 (**0.1622**)	0.2015 (**0.1628**)	0.2044 (**0.1524**)
Accuracy score	0.6863 (**0.7334**)	0.6885 (**0.7553**)	0.7119 (**0.7717**)
Jaccard score	0.6019 (**0.6677**)	0.6038 (**0.6676**)	0.6033 (**0.6839**)

Hamming loss, accuracy score, and Jaccard score are used to evaluate the multi-label model primarily. Hamming loss refers to the average fraction of the wrong prediction of each sublabel. The accuracy score is based on the accuracy of the serial label prediction. Jaccard score measures the proportion of prediction for a sample to its true label. Bold values refer to better performance based on each criterion.

TABLE 5 | Evaluation of partial susceptibility prediction in discovery and replication cohort.

		Precision	Recall		Precision	Recall
Discovery	OSSA	0.81 (**0.88**)	0.85 (**0.90**)	CSSA	0.82 (**0.87**)	0.88 (**0.93**)
Linkou	ORSA	0.86 (**0.91**)	0.81 (**0.89**)	CRSA	0.82 (**0.89**)	0.75 (**0.82**)
Cohort	Accuracy		0.83 (**0.89**)			0.82 (**0.88**)
Replication	OSSA	0.78 (**0.82**)	0.83 (**0.88**)	CSSA	0.83 (**0.84**)	0.86 (**0.90**)
Kaohsiung	ORSA	0.80 (**0.86**)	0.74 (**0.79**)	CRSA	0.75 (**0.82**)	0.71 (**0.72**)
Cohort	Accuracy		0.79 (**0.84**)			0.80 (**0.83**)

The data outside the bracket are generated by the Lowest Power set ensemble with logistic regression. The data within the bracket are generated by the Lowest Power set ensemble with XGBoost. Bold values refer to better performance based on each criterion.

RESULTS

Performance of Multi-Label Prediction Learning Using Logistic Regression and XGBoost

To realize the goal of serial antibiotic resistances prediction, the study adopted three multi-learning ensembles provided by scikit-multilearn 0.2.0, including BinaryRelevance, ClassifierChain, and Lowest Power set. For each ensemble, the study applied LR and XGBoost as the classifier, respectively. Based on the ensembles and classifiers, the prediction model could provide prediction to the susceptibilities of oxacillin and clindamycin within one-step training among the Linkou cohort for each sample. The primary model evaluation among the Kaohsiung cohort is shown in **Table 4**.

From the evaluation criteria shown in **Table 4**, when applying XGBoost as the classifier in all of three multi-label ensembles, the performance in serial label prediction (Accuracy score) or partial label within the prediction (Hamming loss and Jaccard score) both indicated an improved model than using LR as the classifier. Based on the accuracy score, approximately 6% of improvement using XGBoost could be observed from 0.69 on average to 0.75 on average, which presents a refinement that exists in multi-label prediction for antibiotic susceptibility.

Besides the evaluation for the multi-label prediction, analysis that was related to partial or single susceptibility is conducted by dividing the serial label prediction for each sample into susceptibility prediction for oxacillin and clindamycin correspondingly. Herein, the ensemble Lowest Power set with LR and XGBoost is adopted to evaluate the partial performance for its best serial performance in **Table 4**. The evaluation information

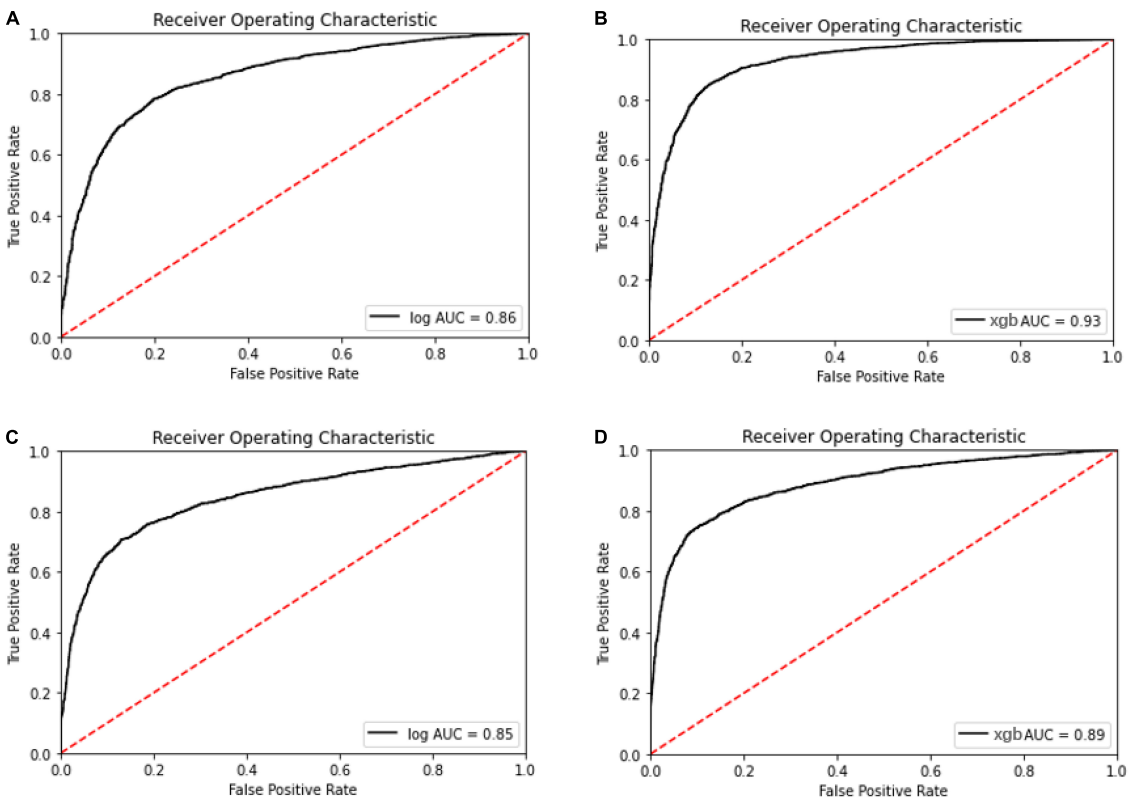


FIGURE 3 | Receiver operating characteristic curve for the oxacillin and clindamycin susceptibility prediction under replication cohort in multi-label learning model, respectively. The area under the curve (AUC) is noted in the curve. **(A)** Receiver operating characteristic (ROC) for the oxacillin susceptibility prediction in replication cohort using Lowest Power set with logistic regression (LR); **(B)** ROC for the oxacillin susceptibility prediction in replication cohort using Lowest Power set with XGBoost; **(C)** ROC for the clindamycin susceptibility prediction in replication cohort using Lowest Power set with LR; and **(D)** ROC for the clindamycin susceptibility prediction in replication cohort using Lowest Power set with XGBoost.

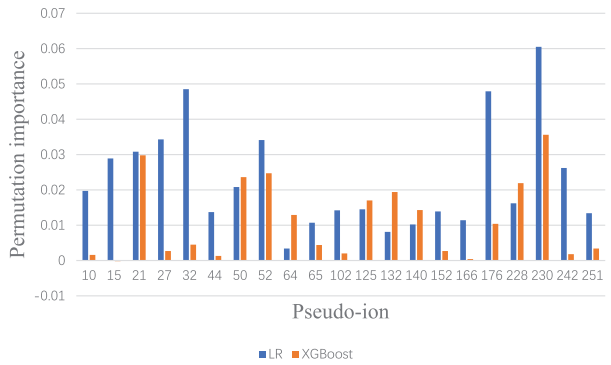


FIGURE 4 | Significant permutation importance comparison between LR and XGBoost under Lowest Power set ensemble.

including precision and recall in each class and total accuracy among discovery and replication cohort is presented in **Table 5**. Both from the discovery and replication cohort, the ensemble that is applied with XGBoost shows more reliable performance than the ensemble that is applied with LR in oxacillin and clindamycin. The discovery accuracy gets increased from 0.83 (oxacillin) and 0.82 (clindamycin) to 0.89 and 0.88. The replication accuracy gets increased from 0.79 (oxacillin) and 0.80 (clindamycin) to 0.84 and 0.83.

Meanwhile, with the consideration of specificity and sensitivity of prediction models, the area under the ROC curve was used to measure the model performance. The AUC of oxacillin and clindamycin prediction using LR or XGBoost as the classifier in the multi-label learning model is shown in **Figure 3**. The AUC for oxacillin susceptibility prediction under XGBoost is 0.93, while it is only 0.86 for the model applying LR. The performance for oxacillin resistance prediction got improved compared with the AUC of 0.80 from DRIAMS (Weis et al., 2022) as well. Meanwhile, the AUC of clindamycin susceptibility prediction gets increased from 0.85 to 0.89 by turning LR into XGBoost.

To visualize and identify the performance improvement, permutation importance of features under Lowest Power set using LR and XGBoost is conducted, respectively. After calculating the permutation importance in each model, each

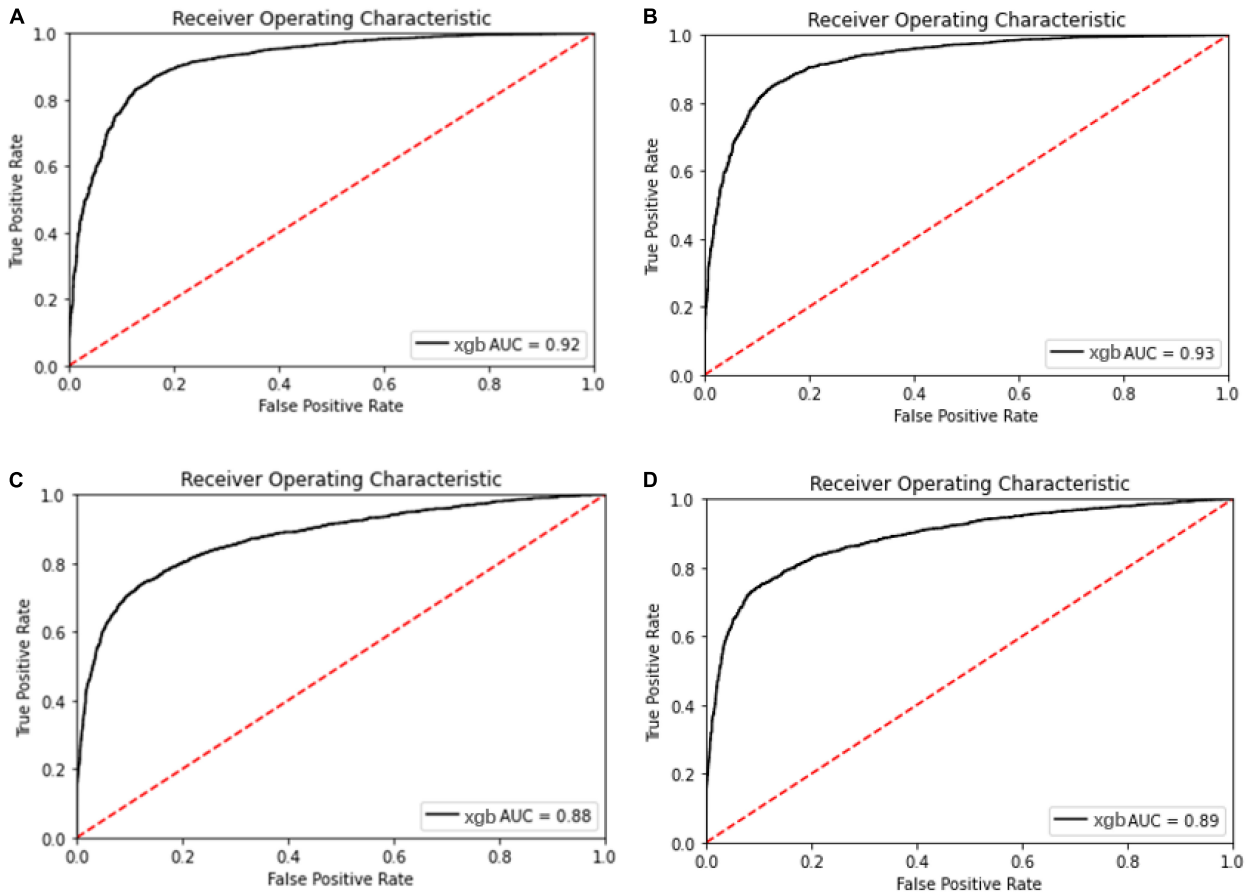


FIGURE 5 | Receiver operating characteristic curve for the oxacillin and clindamycin susceptibility prediction under replication cohort in multi-label learning model, respectively. The AUC is noted in the curve. **(A)** ROC for the oxacillin susceptibility prediction in replication cohort using XGBoost; **(B)** ROC for the oxacillin susceptibility prediction in replication cohort using Lowest Power set with XGBoost; **(C)** ROC for the clindamycin susceptibility prediction in replication cohort using XGBoost; and **(D)** ROC for the clindamycin susceptibility prediction in replication cohort using L Lowest Power set with XGBoost.

feature is assigned with an importance value ranging from 1 to −1, which indicates the feature contribution to the model performance. The whole permutation importance is shown in the **Supplementary Material**. To analyze the main feature importance between LR and XGBoost under Lowest Power set, features that contribute higher than 0.01 permutation importance in either LR or XGBoost model are extracted for comparison (**Figure 4**). Based on permutation importance comparison, the multi-label prediction model using Lowest Power set with LR mainly focuses on pseudo-ion with low and high *m*/*z*-ratio-relatively-in-the-MALDI-TOF-data-(pseudo-ions 10, 15, 27, 32, 44, 176, 230, and 242). For instance, the multi-label model using LR assigns high permutation importance to pseudo ion 15, which stands for protein fragments from 2,310 to 2,330 *m*/*z*, and pseudo ion 242, which stands for protein fragments from 6,850 to 6,870 *m*/*z*. Although the multi-label prediction model using XGBoost shown in **Figure 4** indicates some shared important features, such as pseudo ions 21 and 230, it presents a focus on the medial features, such as pseudo ions 64 and 132. On the whole, the over-refinement between applying LR or XGBoost as the classier in the multi-label prediction model reflects on the relief of permutation importance in the protein fragments with low or high *m*/*z* ratio and a new focus on the medial pseudo ions.

Performance of Multi-Label Prediction Learning and Single Label Prediction Learning

With the hypothesis that whether the model is trained under a single label or multi labels has an influence on the prediction performance, this study constructs two models for oxacillin susceptibility and clindamycin susceptibility, respectively, and separately using XGBoost. Precisely, it refers to using only one susceptibility label from the discovery cohort to train the XGBoost model for the prediction in the replication cohort. During the model training process, RandomOverSampler is adopted to balance the class size from the Python package, Imbalanced-learn (Lemaître et al., 2017). The multi-label prediction model, in this study, still applied XGBoost in the Lowest Power set ensemble.

The ROC and AUC for performance comparison among the multi-label prediction model and single label prediction model are shown in **Figure 5**. Initially, considering due to the multi-label learning, the boundary conditions or hyperplanes for the model may need to be relaxed relatively compared with the single label prediction. However, the model performance using XGBoost as the classier in the multi-label prediction model actually presents approaches to the single susceptibility prediction model and even better to the single one.

To visualize the refinement reflection, permutation importance provides insights to model construction. The full permutation importance of pseudo ion is attached to the **Supplementary Document**. The pseudo ions with permutation importance larger than 0.01 were used for comparison (**Figure 6**). For the oxacillin susceptibility prediction, within the multi-label prediction model, it shows a diverse focus among pseudo ions with different *m*/*z* ratios. For instance, the model assigned more

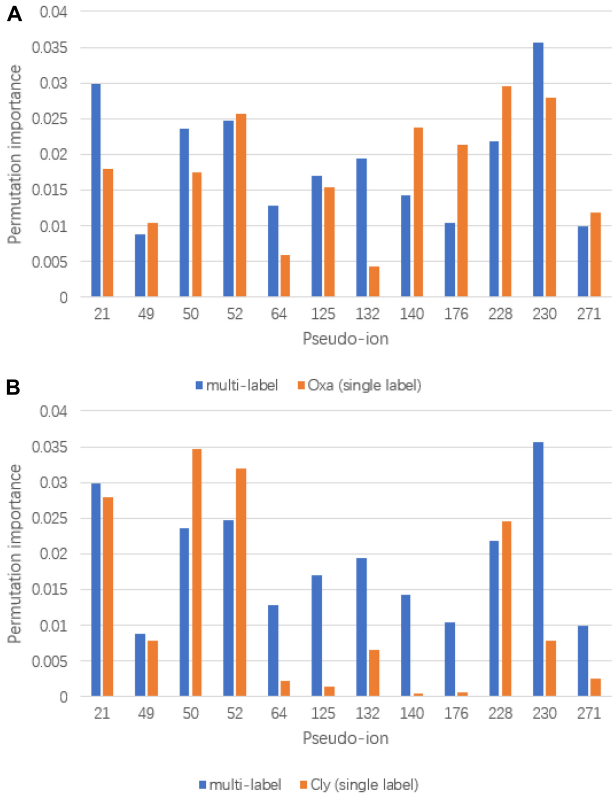


FIGURE 6 | Significant permutation importance comparison between multi-label learning model and single label learning model. **(A)** Permutation importance comparison for oxacillin susceptibility prediction and **(B)** permutation importance comparison for clindamycin susceptibility prediction.

significant permutation importance to pseudo ion 21 with its *m*/*z* ratio from 2,410 to 2,430, ion 50 with its *m*/*z* ratio from 3,010 to 3,030, ion 64 with its *m*/*z* ratio from 3,290 to 3,310, ion 132 with its *m*/*z* ratio from 4,650 to 4,670 and ion 230 with its *m*/*z* ratio from 6,610 to 6,630. In terms of clindamycin susceptibility prediction, the multi-label prediction model shows a consensus that it majorly focuses on pseudo ion with the *m*/*z* ratio from 3,000 to 4,000, determining higher permutation importance on pseudo ions 64, 125, 132, 140, and 176 than the single label model. Besides, compared with the single label prediction model for oxacillin and clindamycin susceptibility together, the multi-label learning model addresses pseudo ions 21, 64, 125, 132, and 230.

DISCUSSION AND CONCLUSION

To deal with potential consideration for the efficiency and accuracy of detection, efforts during the whole experiment are conducted for the realization of a practical model. First, in our study, five kinds of specimens with oxacillin and clindamycin susceptibility labels were used to convince the model. In the future, more kinds of specimens could be used to strengthen our model. Second, the large size of the discovery cohort and

the offer of replication cohort largely support the machine learning model. The size of 26,852 for the discovery cohort from 2013 to 2019 from the Linkou, Taiwan area, is fundamental to train the model solidly. Meanwhile, the replication cohort with 4,963 samples from Kaohsiung served as the validation set. Third, instead of constructing a machine learning model for only one antibiotic susceptibility, our study uses the Lowest Power set ensemble and applies XGBoost as the classifier to build up a multi-label prediction model, which could predict the susceptibilities of the oxacillin and clindamycin at the same time with only one-step training. From the model performance, the multi-label model combined with XGBoost shows better performance (AUC and accuracy) than choosing LR as the classifier, which is commonly used for susceptibility prediction in previous studies (Moradigaravand et al., 2018; Wang et al., 2020). In terms of the output type, the multi-label prediction model performs better than the single label model with only one training process. Furthermore, feature importance was used to analyze the improvement between models, and several potential biological insights were generated.

Based on the prediction performances between single label prediction model and multi-label prediction model and permutation importance results, feature contribution analysis was conducted. For the oxacillin susceptibility prediction, the dominant importance increase of pseudo ions 21 and 50, referring to 2,410–2,430 Da and 2,990–3,010 Da, respectively, matches with the research of Josten et al. (2014). Their study regarded the fragment with an m/z ratio of 2,413 Da as a marker for the presence of phenol-soluble modulin (PSM)-mec, which is a small excreted peptide encoded by the mec gene. Besides, the PSM-mec is excreted by agr-positive strains, where it presents with the delta-toxin with an m/z ratio of 3,007 Da. For the clindamycin susceptibility prediction, fragment around pseudo ion 64, the m/z ratio of 3,270 Da–3,290 Da is present to show the expression of Cfr. Gene Cfr induces the resistance to clindamycin. Beyond pseudo ion 64, ions 50 and 52 could be potential entry for biological insight analysis for their high permutation importance on clindamycin susceptibility prediction. The synergy importance increase occurred on pseudo ions 64, 132, and 230, covering m/z of 3,270–3,290 Da, 4,630–4,650 Da, and 6,590–6,610 Da. The hypothesis of synergy effect between the susceptibility of oxacillin and clindamycin or even among multidrug resistance could likely be tested by considering the three ions above.

There are several limitations and restrictions in our study. First, the discovery and replication cohorts are actually based on the local part in the Taiwan area, and the model needs more samples worldwide to become a practical susceptibility prediction model at a global level. Second, in our study, we only possessed and considered the susceptibilities of oxacillin and clindamycin, the simplest case of the multi-label prediction. There is a simple correlation analysis between two susceptibilities. In the future, during the sample recruitment, information about more than three kinds of susceptibilities could be tested and collected to realize more complex serial label predictions. Meanwhile, the statistical methods for isolate selection for MALDI-TOF need to be improved. Previous study has concluded that single isolate selection in MALDI-TOF may generate biased results if missed to identify the diversity among isolates (Pinar-Méndez et al., 2021). Considering the variation among isolates, the MALDI-TOF result from one isolate for each patient is not representative enough as the input data for susceptibility prediction. Optimized statistical methods are needed, such as multi-isolate selection or MS data integration from multi isolates, which could be conducted in the future study for a comprehensive prediction model. In addition, future studies can adopt a novel ensemble method that considers the relation among labels or susceptibilities instead of Lowest Power set in this study for better serial prediction performance. A platform or database combining resistance information, such as prediction or tendency with large sample size and diverse drug susceptibilities, could be continuous for future study. During the permutation importance analysis, some pseudo ions were pointed out to respond for the model refinement. These ions could be considered as the potential biomarker or functional segments and needed to analyze in the laboratory. Although the AUC for oxacillin and clindamycin susceptibility prediction indicates good performances, the accuracy for serial susceptibility prediction still does not satisfy the clinical requirements. However, our model presents the possibility of a proteomic-based model with a machine learning algorithm for rapid serial susceptibility prediction.

To summarize our study, we successfully constructed a multi-label prediction model applying XGBoost in Lowest Power set for oxacillin and clindamycin susceptibilities based on the MALDI-TOF MS data with the output of serial labels. Multidrug resistance is a threat to disturb treatment effects and usually tested by AST, which is limited by the labor and facility resource. Under large-scale size in the discovery cohort and replication cohort, our model could realize serial susceptibility prediction solidly, which ideally help patients and doctor with clinical guidance and insights to the antibiotic usage efficiently and accurately. In a nutshell, combing MALDI-TOF MS and machine learning algorithm will widely spread a proteomic-based antibiotic susceptibility test clinically taking advantage of speed and accuracy and saving the resources that originally are consumed for the costing and inefficient AST.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

JZ, ZW, H-YW, and T-YL conceived the project, designed and conducted the analyses, interpreted the results, wrote the manuscript, and are listed in random order. H-YW collected the clinical samples and executed the microbiology experiments. JZ and ZW conducted the analyses and wrote the manuscript. JZ assisted with the machine learning analysis and visualization. C-RC and J-TH assisted with manuscript revision.

J-JL and T-YL supervised the study. All authors have read and approved the manuscript.

FUNDING

This work was supported by the Guangdong Province Basic and Applied Basic Research Fund (2021A1515012447), National Natural Science Foundation of China (32070659), the Science, Technology and Innovation Commission of Shenzhen Municipality (JCYJ20200109150003938), Ganghong Young Scholar Development Fund (2021E007), ShenzhenHong Kong Cooperation Zone for Technology and Innovation (HZQB-KCZYB-2020056), Chang Gung Memorial Hospital (CMRPG3L0401 and CMRPG3L0431), and the Ministry of Science and Technology, Taiwan (111-2320-B-182A-002 -MY2). This work was also supported by the Warshel Institute for Computational Biology, School of Life and Health Sciences, The Chinese University of Hong Kong, Shenzhen, China.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.853775/full#supplementary-material

REFERENCES

Altmann, A., Toloşi, L., Sander, O., and Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics* 26, 1340–1347. doi: 10.1093/bioinformatics/btq134

Bell, J. M., and Turnidge, J. D. (2002). High prevalence of oxacillin-resistant *Staphylococcus aureus* isolates from hospitalized patients in Asia-pacific and South Africa: results from sentry antimicrobial surveillance program, 1998–1999. *Antimicrob. Agents Chemother.* 46, 879–881. doi: 10.1128/AAC.46.3.880-882.2002

Chen, T., and Guestrin, C. (2016). "Xgboost: a scalable tree boosting system," in *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining* (New York, NY: ACM), 785–794.

Croxatto, A., Prod'hom, G., and Greub, G. (2012). Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiol. Rev.* 36, 380–407. doi: 10.1111/j.1574-6976.2011.00298.x

Hiramatsu, K., Katayama, Y., Matsuo, M., Sasaki, T., Morimoto, Y., Sekiguchi, A., et al. (2014). Multi-drug-resistant *Staphylococcus aureus* and future chemotherapy. *J. Infect. Chemother.* 20, 593–601. doi: 10.1016/j.jiac.2014.08.001

Hodille, E., Badiou, C., Bouveyron, C., Bes, M., Tristan, A., Vandenesch, F., et al. (2018). Clindamycin suppresses virulence expression in inducible clindamycin-resistant *Staphylococcus aureus* strains. *Ann. Clin. Microbiol. Antimicrob.* 17:38. doi: 10.1186/s12941-018-0291-8

Josten, M., Dischinger, J., Szekat, C., Reif, M., Al-Sabti, N., Sahl, H. G., et al. (2014). Identification of agr-positive methicillin-resistant *Staphylococcus aureus* harbouring the class A mec complex by MALDI-TOF mass spectrometry. *Int. J. Med. Microbiol.* 304, 1018–1023. doi: 10.1016/j.ijmm.2014.07.005

Kehrenberg, C., Schwarz, S., Jacobsen, L., Hansen, L. H., and Vester, B. (2005). A new mechanism for chloramphenicol, florfenicol and clindamycin resistance: methylation of 23S ribosomal RNA at A2503. *Mol. Microbiol.* 57, 1064–1073. doi: 10.1111/j.1365-2958.2005.04754.x

Lay, J. O. (2001). MALDI-TOF mass spectrometry of bacteria. *Mass Spectrom. Rev.* 20, 172–194. doi: 10.1002/mas.10003

Lemaître, G., Nogueira, F., and Aridas, C. K. (2017). Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning. *J. Mach. Learn. Res.* 18, 559–563.

Louie, L., Matsumura, S. O., Choi, E., Louie, M., and Simor, A. E. (2000). Evaluation of three rapid methods for detection of methicillin resistance in *Staphylococcus aureus*. *J. Clin. Microbiol.* 38, 2170–2173. doi: 10.1128/jcm.38.6.2170-2173.2000

Moradigaravand, D., Palm, M., Farewell, A., Mustonen, V., Warringer, J., and Parts, L. (2018). Prediction of antibiotic resistance in *Escherichia coli* from large-scale pan-genome data. *PLoS Comput. Biol.* 14:e1006258. doi: 10.1371/journal.pcbi

Pinar-Méndez, A., Fernández, S., Baquero, D., Vilaró, C., Galofré, B., González, S., et al. (2021). Rapid and improved identification of drinking water bacteria using the drinking water library, a dedicated MALDI-TOF MS database. *Water Res.* 203:117543. doi: 10.1016/j.watres.2021.117543

Reygaert, W. C. (2013). "Antimicrobial resistance mechanisms of Staphylococcus aureus," in *Microbial Pathogens and Strategies for Combating Them: Science, Technology and Education*, ed. A. Nendez-Vilas (Spain: Formatex Research Center), 297–305.

Schwarz, S., Silley, P., Simjee, S., Woodford, N., van Duijkeren, E., Johnson, A. P., et al. (2010). Editorial: assessing the antimicrobial susceptibility of bacteria obtained from animals. *J. Antimicrob. Chemother.* 65, 601–604. doi: 10.1093/jac/dkq037

Swenson, J. M., Williams, P. P., Killgore, G., O'Hara, C. M., and Tenover, F. C. (2001). Performance of eight methods, including two new rapid methods, for detection of oxacillin resistance in a challenge set of *Staphylococcus aureu* s organisms. *J. Clin. Microbiol.* 39, 3785–3788. doi: 10.1128/JCM.39.10.3785-3788.2001

Szymański, P., and Kajdanowicz, T. (2017). A scikit-based python environment for performing multi-label classification. *arXiv [Preprint]* arXiv:1702.01460,

Wang, Z., Wang, H. Y., Chung, C. R., Horng, J. T., Lu, J. J., and Lee, T. Y. (2020). Large-scale mass spectrometry data combined with demographics analysis rapidly predicts methicillin resistance in *Staphylococcus aureus*. *Brief. Bioinform.* 22:bbaa293. doi: 10.1093/bib/bbaa293

Weis, C., Cuénod, A., Rieck, B., Dubuis, O., Graf, S., Lang, C., et al. (2022). Direct antimicrobial resistance prediction from clinical MALDI-TOF mass spectra using machine learning. *Nat. Med.* 28, 164–174. doi: 10.1038/s41591-021-01619-9

Zhang, M. L., and Zhou, Z. H. (2014). A review on multi-label learning algorithms. *IEEE Trans. Knowl. Data Eng.* 26, 1819–1837. doi: 10.1109/tkde.2013.39

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 06 June 2022
doi: 10.3389/fmicb.2022.821233

Efficiently Predicting Vancomycin Resistance of *Enterococcus Faecium* From MALDI-TOF MS Spectra Using a Deep Learning-Based Approach

Hsin-Yao Wang [1,2†], Tsung-Ting Hsieh [3†], Chia-Ru Chung [4], Hung-Ching Chang [3], Jorng-Tzong Horng [1,4,5], Jang-Jih Lu [1,6,7*] and Jia-Hsin Huang [3*]*

[1] Department of Laboratory Medicine, Chang Gung Memorial Hospital, Taoyuan, Taiwan, [2] Ph.D. Program in Biomedical Engineering, Chang Gung University, Taoyuan, Taiwan, [3] Taiwan AI Labs, Taipei, Taiwan, [4] Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan, [5] Department of Bioinformatics and Medical Engineering, Asia University, Taichung, Taiwan, [6] School of Medicine, Chang Gung University, Taoyuan, Taiwan, [7] Department of Medical Biotechnology and Laboratory Science, Chang Gung University, Taoyuan, Taiwan

Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) has recently become a useful analytical approach for microbial identification. The presence and absence of specific peaks on MS spectra are commonly used to identify the bacterial species and predict antibiotic-resistant strains. However, the conventional approach using few single peaks would result in insufficient prediction power without using complete information of whole MS spectra. In the past few years, machine learning algorithms have been successfully applied to analyze the MALDI-TOF MS peaks pattern for rapid strain typing. In this study, we developed a convolutional neural network (CNN) method to deal with the complete information of MALDI-TOF MS spectra for detecting *Enterococcus faecium*, which is one of the leading pathogens in the world. We developed a CNN model to rapidly and accurately predict vancomycin-resistant *Enterococcus faecium* (VRE*fm*) samples from the whole mass spectra profiles of clinical samples. The CNN models demonstrated good classification performances with the average area under the receiver operating characteristic curve (AUROC) of 0.887 when using external validation data independently. Additionally, we employed the score-class activation mapping (CAM) method to identify the important features of our CNN models and found some discriminative signals that can substantially contribute to detecting the ion of resistance. This study not only utilized the complete information of MALTI-TOF MS data directly but also provided a practical means for rapid detection of VRE*fm* using a deep learning algorithm.

Keywords: vancomycin-resistant *Enterococcus faecium* (VRE*fm*), antibacterial drug resistance, MALDI-TOF MS, convolutional neural network (CNN), rapid detection

OPEN ACCESS

Edited by:
Karsten Becker, University Medicine Greifswald, Germany

Reviewed by:
Dana Marshall, Meharry Medical College, United States
Nadjette Bourafa, Mohamed-Cherif Messaadia University - Souk Ahras, Algeria

***Correspondence:**
Jorng-Tzong Horng
horng@db.csie.ncu.edu.tw
Jang-Jih Lu
janglu45@gmail.com
Jia-Hsin Huang
jiahsin.huang@ailabs.tw

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

Received: *24 November 2021*
Accepted: *28 April 2022*
Published: *06 June 2022*

Citation:
Wang H-Y, Hsieh T-T, Chung C-R, Chang H-C, Horng J-T, Lu J-J and Huang J-H (2022) Efficiently Predicting Vancomycin Resistance of Enterococcus Faecium From MALDI-TOF MS Spectra Using a Deep Learning-Based Approach. Front. Microbiol. 13:821233. doi: 10.3389/fmicb.2022.821233

INTRODUCTION

Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) has become a promising analytical technique in many clinical microbiology laboratories in identifying bacterial species. However, applying MALDI-TOF MS in determining antibiotic susceptibility test (AST) has not been widely developed. The traditional approach utilized the presence or absence of several peaks on MS spectra to predict the AST (Wolters et al., 2011; Lasch et al., 2014). However, the predictive performance based on the traditional approach is discrepant from each other, hindering the application in clinical settings. The discrepancy in predictive performance would have resulted from the insufficient number of peaks used in these studies. Moreover, single peaks' presence or absence rather than the whole pattern of the peaks were used for classifying AST. In the past few years, some studies have harnessed artificial intelligence (AI) algorithms to analyze the MALDI-TOF MS peaks pattern for classifying specific bacterial strains (Wang et al., 2019, 2020b; Weis et al., 2022). Most of the works studied *Staphylococcus aureus* (Weis et al., 2022), group B *Streptococcus* (Wang et al., 2019), and Enterobacteriaceae (Weis et al., 2022). By contrast, *Enterococcus faecium* is also a superbug with rising clinical importance (Ahmed and Baptiste, 2018), only a few studies have been reported for rapid detection of vancomycin-resistant *Enterococcus faecium* (VRE*fm*) by using MALDI-TOF MS and machine learning approaches (Griffin et al., 2012; Wang et al., 2021). Rapid detection of VRE*fm* would result in favorable clinical outcomes, including reduced mortality rate and reduced use of broad-spectrum antibiotics for severe VRE*fm* infection.

A recent benchmark study has demonstrated that a one-dimensional convolutional neural network (CNN) outperformed traditional machine learning methods for the bacterial species identification based on the MALDI-TOF MS data (Mortier et al., 2021). Deep learning methods, such as CNNs and recurrent neural networks (RNNs), have been repeatedly proved to outperform the classical machine-learning algorithms for large datasets with high-dimensional data. A CNN applies filters in the form of a convolution operation to extract features from the data. The advantage of a CNN is that it reduces the parameters compared to other neural networks by sharing them as multiple filters (Lecun et al., 1998). The convolution method, which is coupled with a small grid of input signals into local receptive features, could be a solution to solve the peak shift among samples to apply raw data directly. Furthermore, the convolution filters share the parameters independent of position; thus, we can reduce the number of used parameters. This parameter sharing of the convolutional filters and the local connections of the nodes increases the performance in handling sparsely connected data.

In this study, we aim to propose a CNN algorithm for practical extraction and analysis of multidimensional MS spectral data for VRE*fm* prediction. By using the consecutively collected MALDI-TOF MS data from large clinical isolates from two tertiary medical centers (Chang Gung Memorial Hospital [CGMH], Linkou branch, and Kaohsiung branch) in Taiwan (Wang et al., 2021), we coupled with the respective laboratory-confirmed antibiotic resistance profile to build CNN models for antimicrobial resistance prediction. We first demonstrated the efficacy of our CNN architectures for discrimination of VRE*fm* from vancomycin-susceptible *Enterococcus faecium* (VSE*fm*) isolates in reporting the area under the receiver operating characteristic (AUROC) value over 0.800. Next, a full CNN model trained with data from the Linkou branch achieved good performance with an average AUROC of 0.887 in predicting VRE*fm* in the independent data from the Kaohsiung branch. Finally, we applied the Score-CAM method and statistical analysis to examine the important features that the CNN models used to predict VRE*fm* isolates. Of note, some essential features of the CNN model were reported in the literature and many m/z ranges were novel features, showing significant differences in m/z intensities of the MALDI-TOF MS spectra in the VRE*fm* from clinical susceptible isolates. Furthermore, since mass spectra can be generated rapidly from colonies following an overnight culture, we provide an efficient CNN framework to build a prediction model for antimicrobial resistance based on the complete MALDI-TOF MS profiles directly and, therefore, could help the clinical management of patients with infectious diseases.

MATERIALS AND METHODS

Data Sources

Matrix-assisted laser desorption ionization time-of-flight MS spectra of VRE*fm* were used as the input features, while the susceptibility test to vancomycin was used as the label of interest. MALDI-TOF MS (Bruker Daltonik GmbH, Bremen, Germany) spectra of VRE*fm* were consecutively collected between 2013 and 2017 in Chang Gung Memorial Hospital, Linkou, and Kaohsiung branches. The manufacturer's instruction was followed, and default settings were used to identify *E. faecium* (Wang et al., 2018a). Biotyper 3.1 software (Bruker, Germany) was used for species identification of *E. faecium*. Regarding labeling, the susceptibility to vancomycin was determined by using the paper disc method based on the CLSI M100 guideline (CLSI, 2020). Accordingly, the detailed specimen distribution of VRE*fm* and VSE*fm* clinical isolates is summarized in **Supplementary Figures 1A,B** in the Linkou and Kaohsiung branches, respectively. The MALDI-TOF spectra were preprocessed by using Flexanalysis (Bruker, Germany). The relative intensity threshold was set as zero. The signal-to-noise ratio (S/N ratio) of two was used to filter out signals whose S/N ratio was lower than two. Baseline subtraction of the raw spectra was done by using the top-hat algorithm. Savitzky-Golay algorithm was adopted to smooth the spectra. MS spectra preprocessed by the above methods were used as the input data for the subsequent modeling.

CNN Model

In this study, we used one-dimensional CNNs to classify the *E. faecium* strain with vancomycin resistance (VRE*fm*) according to the raw MS data. Herein, the MS raw data represent the input data and the goal is to predict whether the given sample is VRE*fm* (class label 1) or not (VSE*fm*, class label 0). In total, 7997 *E. faecium* cases were identified and included from the Linkou and Kaohsiung branches of CGMH, while 4,017 cases

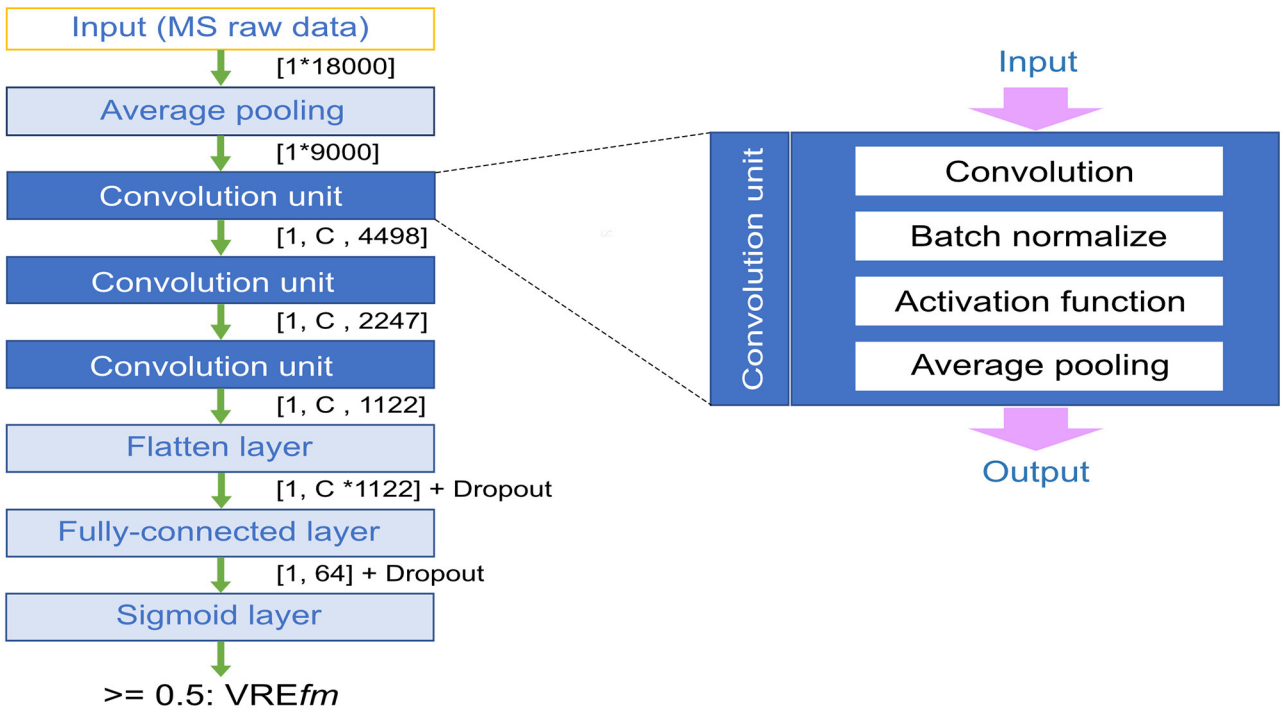


FIGURE 1 | Illustration of convolutional neural network (CNN) model architecture. C denotes different channels, and the dropout rate was 0.5.

were VRE*fm* (50.23%) and 3,980 cases were VSE*fm* (49.77%) cases, respectively.

Figure 1 shows a CNN model architecture used in this study. Because the whole m/z signals of raw MS data represented a vector consisting of 18,000 values, an average pooling (avg pooling) layer is used for the summation of raw signals from the nearby m/z peaks, which are sometimes shifted among samples. A CNN model can learn local m/z signal patterns, which are good discriminators between positive and negative instances in the training dataset. We added three convolutional layers to learn the information across the patterns of m/z peaks. The information of adjacent m/z signals is embedded in the entries of kernels (filters) used in the first convolution layer. Additional CNN layers can learn higher-order interactions between signals at different m/z regions. A convolution layer comprises four units, including convolution filter or kernel, batch normalization, activation function, and pooling. After the convolution, a flattened layer was applied to conjugate the high-level information of m/z peak patterns.

In this study, we tried several versions of CNN models including preserving the average pooling layer (on or off), different activation functions using tanh or rectified linear units (ReLu), and a varying number of channels in the convolution layers (32, 64, 128), and the dropout rate (0 or 0.5) in the final fully connected layer. The output layer contained a single neuron with the Sigmoid activation function, which learns the mapping from the hidden (fully connected) layer to the output class labels [0, 1]. The final output is a probability indicating whether an input is a vancomycin-resistant *E. faecium* strain.

Training and Testing

To develop a robust CNN model, we applied 10-fold cross-validation in the data from the CGMH Linkou branch to evaluate the model performance for the optimal CNN architecture. Next, the whole data from the CGMH Linkou branch were used as the training set to develop the final CNN model; the data from the CGMH Kaohsiung branch served as the unseen independent testing set for the external validation. In making a performance comparison on the testing set, we also applied two commonly used machine learning (ML) algorithms, i.e., random forest (RF) (Breiman, 2001) and extreme gradient boosting (XGBoost) (Chen and Guestrin, 2016), to build the VRE*fm* prediction models. For the ML-based models, the input attributes of the important peaks were selected by the feature selection method used in the previous study (Wang et al., 2021). The primary parameters of the two ML models were summarized in **Supplementary Table 1**. In addition, we developed those models using ten different random seeds for performance evaluation and feature selection.

Performance Measurement of Predictive Models

The VRE*fm* prediction models trained using CNN algorithms were evaluated *via* 10-fold cross-validation (CV) using the data from the Linkou branches of CGMH. In the 10-fold CV, all the data from Linkou branches of CGMH were randomly divided into ten subgroups with approximately equal data sizes. Each fold of subgroups was used as the independent testing data to

evaluate the performance of the model trained on the other dataset. To evaluate the model performance, accuracy (ACC) and the area under the receiver operating characteristic (AUROC) were considered as the primary metrics for the performance comparison. After obtaining the CNN model architecture with better performance, all data from the Linkou branches of CGMH were used to train a full model. The data from the Kaohsiung branch of CGMH served as the unseen independent testing data for external validation, method comparison, and feature selection.

Feature Selection Procedures

We selected the top and last 500 samples from the CGMH Kaohsiung branch data according to the prediction scores from the best CNN models. In order to see where the CNN model learned from the MS peak patterns, we applied the selected samples to verify the important weight of each m/z peak of the CNN models using the Score-CAM technique (Wang et al., 2020a). First, we ranked the Score-CAM scores of each m/z peak, chose the highest 1% of peaks and conjugated the adjacent m/z peaks into a range. Then, the informative features were identified if the m/z peak ranges were selected at least eight times among ten independent CNN models.

We further investigated the peak intensities of the informative features and box plots were adopted to display their distributions. For each m/z peak, the intensity values of all samples from the CGMH Kaohsiung branch were normalized by the *Z*-score transformation. Then, the normalized *Z*-scores of the given m/z peaks within the informative features were averaged for every sample. Of note, for a given informative feature, a few isolates were removed from statistical comparison when the intensity for the feature was zero.

Statistics

Statistical analysis was performed by using R software version 4. The differences of multiple groups were analyzed by one-way ANOVA and followed by the Tukey *post hoc* test. In addition, the differences in the normalized intensity of the informative features between VRE*fm* and VSE*fm* isolates were calculated by Wilcoxon rank-sum test and the significant level was controlled for multiple testing using the *q-value* method (Storey et al., 2020).

Code Availability

The computer codes that support the findings of this study were deposited to the GitHub repository and are available at https://github.com/p568912/CNN_MALDI-TOF_VREfm.

RESULTS

Development of Deep Learning-Based VRE*fm* Prediction Models

We aimed to develop a robust deep learning-based VRE*fm* prediction model using the complex MALDI-TOF m/z spectra without predictive peak selection in the data preprocessing as shown in the previous study (Wang et al., 2021). We used the cases from the CGMH Linkou branch to evaluate the architectures and parameters of the CNN models with a 10-fold cross-validation approach. Next, we compared CNN model performance with the other two machine learning algorithms using the cases from the CGMH Kaohsiung branch as the unseen independent testing dataset. Finally, we trained ten CNN models independently by setting different seeds for model initiation and examining the prediction performance with two matrices including ACC and AUROC.

We first tested different CNN architectures with the pooling techniques on the input data and different activation functions. As the pooling layer provides an approach to summarize the MS peak intensities across the nearby region, the slight shift of MS peaks across samples is expected to be solved. Indeed, the variation of accuracies across 10-fold CV was much reduced when the presence of a pooling layer (**Figure 2A**). When ReLu as the non-linear activation function was used in the neurons, the model performance with a prior pooling layer yielded the best performance in the models (**Figure 2B**).

Next, we tried to decrease or increase the channel sizes to 32 and 128, respectively, for each convolution layer and added dropout, which could avoid the overfitting problem. Without dropout function, the ACC and AUROC of those models with lower (32) or higher (128) channels were generally lower than the models with a channel size of 64 (**Figure 3A**). However, we observed that the proceeding dropout during training could significantly improve the performance in prediction by showing better AUROC values (**Figure 3B**).

Comparison Between Machine Learning Models and CNN Models

The previous study applied a decision tree-based algorithm named RF to predict the VRE*fm* based on the selected m/z peaks from manual alignment and statistical tests (Wang et al., 2021). Accordingly, we constructed RF-based classifiers and XGBoost-based classifiers with the selected m/z peaks and compared them to our CNN models using whole MS spectra. In this study, the MALDI-TOF MS spectra obtained from the CGMH Kaohsiung branch were regarded as the independent testing dataset. To evaluate the performance fairly, the training dataset was from the Linkou branch and the cases from the Kaohsiung branch was set as unseen testing dataset. The results showed in **Figure 4**. CNN models attained higher performance in predicting VRE*fm* (average accuracy = 0.796) than RF and XGBoost (average ACC = 0.762 and 0.772, respectively). Furthermore, the average AUROC of the CNN algorithm achieved a better performance of 0.887 than the other RF and XGBoost algorithms (average AUROC = 0.845 and 0.855, respectively) (**Figure 4B**).

CNN Models Capture Important Features to Predict VRE*fm*

We applied the Score-CAM (Wang et al., 2020a) to examine the importance of m/z signals as informative features to affect the prediction performance of the best model for the top 500 VRE*fm* and VSE*fm* prediction scores, respectively. Score-CAM was introduced to visually explain how the CNN models classified the MALDI-TOF MS signal patterns into two groups. According

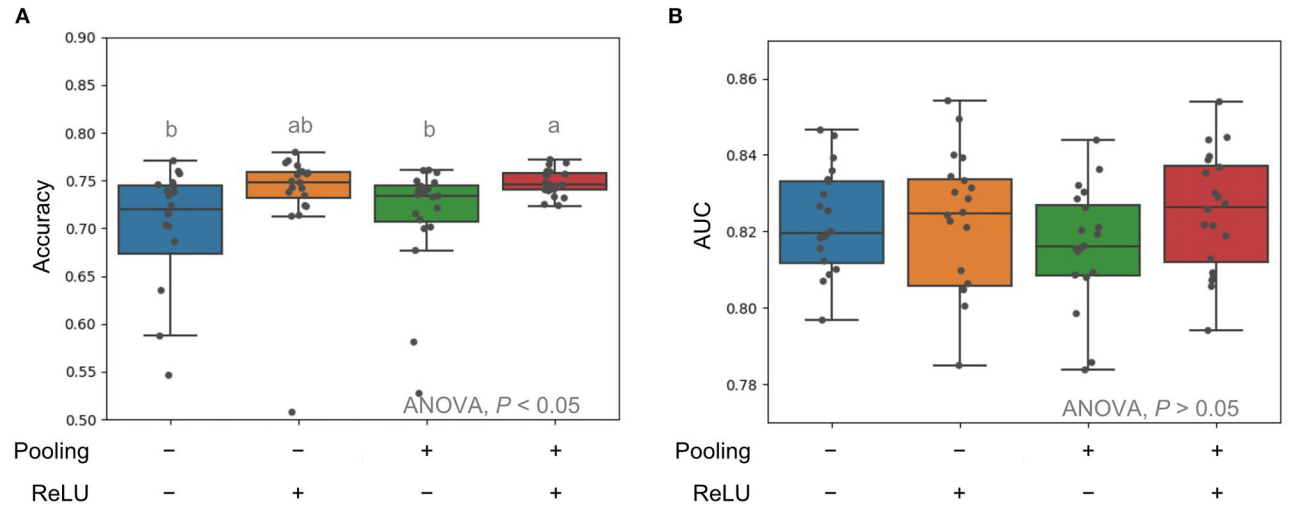


FIGURE 2 | Comparison of average pooling in terms of accuracy **(A)** and AUC **(B)** for input mass spectrometry (MS) raw data and rectified linear units (ReLu) as activation function. Multiple comparison tests on the different units were applied in the models using one-way ANOVA, following the *post-hoc* Tukey honest significant difference (HSD) test for paired difference with a p-value < 0.05. Different letters indicate a significant difference.

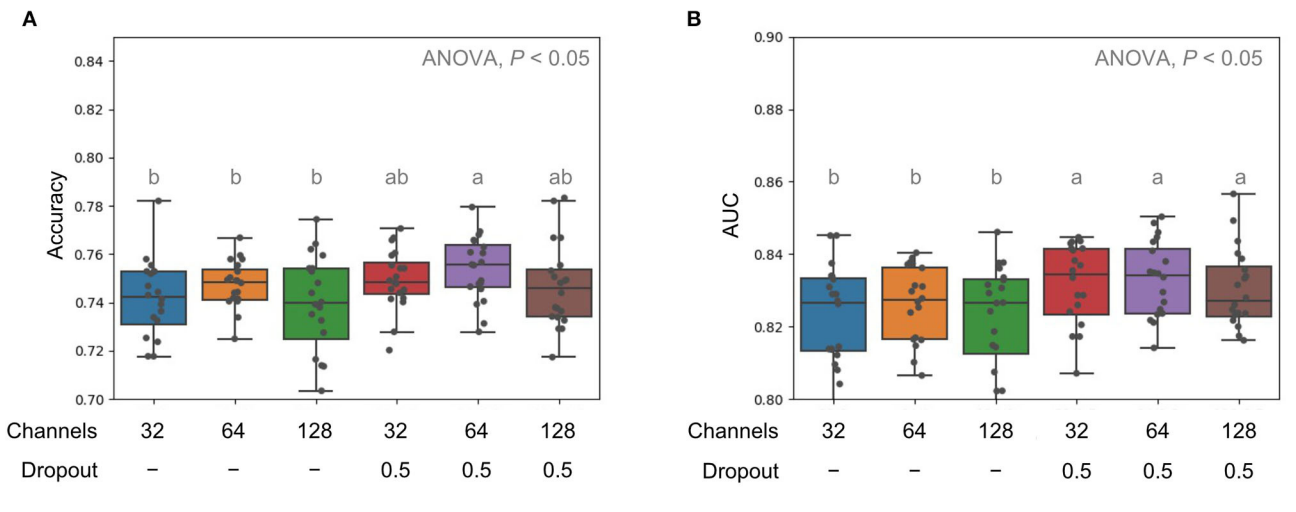


FIGURE 3 | Comparison of model performance in terms of accuracy **(A)** and AUC **(B)** for with different channel sizes and dropouts. Multiple comparison tests on the different units were applied in the models using one-way ANOVA, following the *post-hoc* Tukey HSD test for paired difference with a p-value < 0.05. Different letters indicate a significant difference.

to the Score-CAM scores and selection procedures (see Materials and Methods), the informative features as the predictive m/z peak ranges for making the final categorization decision as VRE*fm* or VSE*fm* were shown in **Figures 5A,B**, respectively. First, the informative features for classifying respective VRE*fm* and VSE*fm* were mostly overlapped. That is, the CNN models could evaluate the patterns of the mass peak ranges to distinguish VRE*fm* and VSE*fm* consistently. Second, many m/z peak signatures, have been reported to identify the vancomycin-resistant *E. faecium*. in the previous works of literature (Lasch et al., 2014; Wei et al., 2014; Wang et al., 2021), were included in the ranges of our informative features. For example, six out of ten most critical predictive peaks for VRE*fm* prediction in our previous work (Wang et al., 2021) have been identified as the important feature ranges of the CNN models. Third, several informative features were specific to classify VRE*fm* such as 3,301–3,304 Da, 5,114–5,118 Da, 5,197–5,200 Da, 5,247–5,253 Da, and 6,602–6,608 Da. In comparison with the statistical method to extract crucial peaks in our previous study (Wang et al., 2021), the occurrence of two peaks of m/z 3,302 and m/z 6,603 were concordantly found in the current CNN model.

We selected the 30 informative features based on the Score-CAM results and further illustrated the differences in intensities of spectra between VRE*fm* and VSE*fm* (**Figure 5C**). Notably,

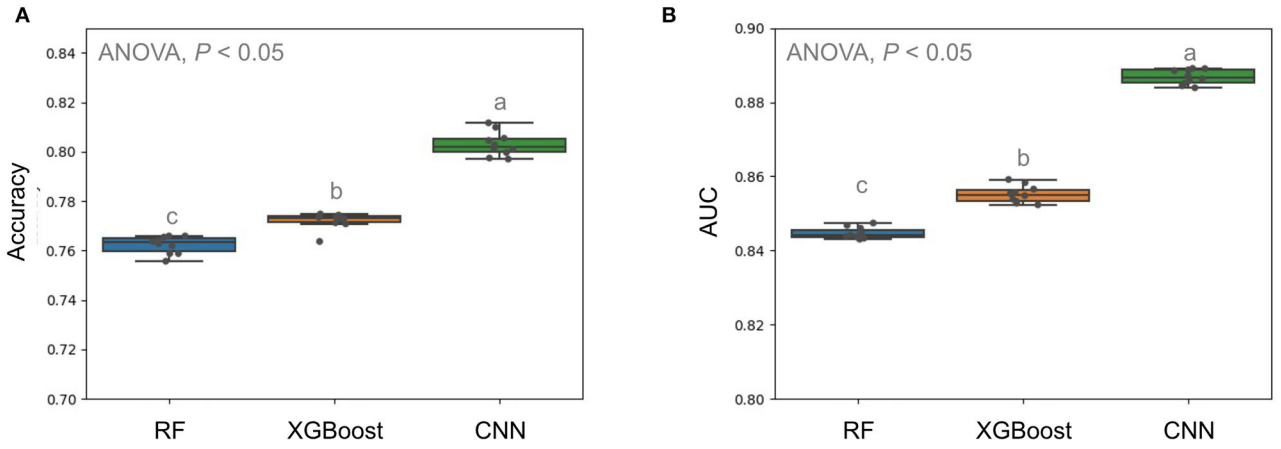


FIGURE 4 | Comparison of prediction performance in terms of accuracy **(A)** and AUC **(B)** using different models. Multiple comparison tests on the different units were applied in the models using one-way ANOVA, following the *post-hoc* Tukey HSD test for paired difference with a p-value < 0.05. Different letters indicate a significant difference.

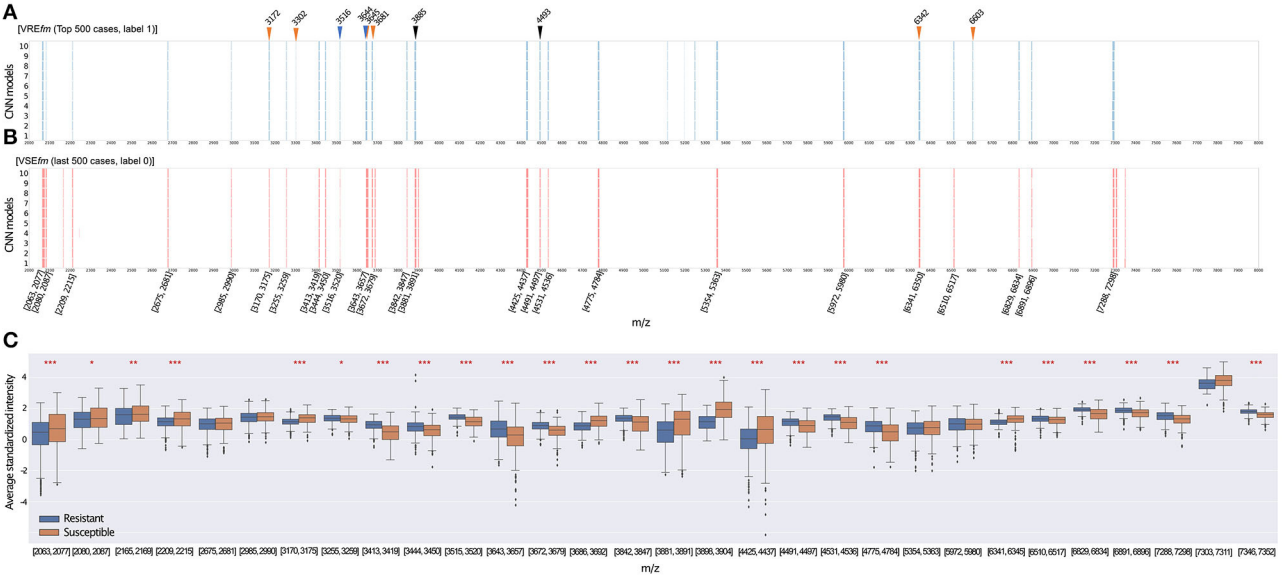


FIGURE 5 | Feature selection is based on the Score-CAM. Each row of m/z peaks represents the important features with top 1% weights in the different models. The important features of our CNN models for the classification of vancomycin-resistant *E. faecium* (VRE*fm*) **(A)** and vancomycin-susceptible *E. faecium* (VSE*fm*) **(B)** in the last and top 500 testing cases, respectively. Rectangles represent the important features of m/z ranges selected by at least 8 independent CNN models using Score-CAM. Triangles indicate that the important markers at m/z signals for discrimination of VRE*fm* strain have been identified in the literature. Blue is identified by Wei et al. (2014), black is identified by Lasch et al. (2014), and orange is identified by Wang et al. (2020b). **(C)** The intensity distribution of informative features with the normalized intensities of 500 resistant and 500 susceptible isolates. The star (*) indicates a statistical difference in the feature intensities between resistant and susceptible isolates. Wilcoxon-rank sum test was applied to test the difference between the two groups. * q-value < 0.05, ** q-value < 0.01, ***, q-value < 0.001.

25 out of 30 informative features in terms of their average m/z peak intensities showed a significant difference between the resistant and susceptible isolates (q-value < 0.05, Wilcoxon rank-sum test). The important features in the Score-CAM were apparently in accordance with the observed results of the MALDI-TOF signal profiles (**Supplementary Figure 3**). This investigation implied that the Score-CAM method successfully captured the critical signals to distinguish the VRE*fm* and VSE*fm* isolates by the CNN model. More importantly, the results provide an explanation that our CNN models could automatically detect the important features to obtain high performance reasonably.

DISCUSSION

This study focused on the classification of resistance strain typing of *E. faecium* based on the MALDI-TOF MS utilizing CNN algorithm method directly. Specifically, the complete raw spectra data did not need to undergo the time-consuming preprocessing steps such as peak alignment of individual samples and statistical analysis to filter out crucial peaks. Instead, we utilized the average pooling method to solve the problem of peak shift across samples and the deep learning technique to learn the complex interaction between m/z peaks related to the resistant and susceptible *E. faecium* strain typing. The rapid identification of VRE*fm* strain types will facilitate the speed in the identification of suspicious infections and provide suitable treatment to patients for rapid infection control. Additionally, we further explore the discriminative features in the CNN model, which was considered a black box, and will allow the identification of each corresponding peptide with further biological experiments as the followings. Such findings should provide clinically valuable information pertaining to the different subtypes of *E. faecium* and other resistant bacterial strains.

In order to handle the issue of peak shifting, alternative approaches were used including "type templates" for each susceptible type based on the incidence of specific peaks in their MALDI-TOF MS spectra (Wang et al., 2018a,b), recursive search based on the statistical analysis and performance of classifiers (Chung et al., 2019) and the embedded feature-selection method (Wang et al., 2021). Different approaches were designed to deal with the peak shift problem of using spectral data with particular procedures. In contrast, we here use the average pooling, which is a straightforward approach, to consider the peak shift problem to tolerant the difference across the spectra data. Although the average pooling might compromise the crucial peak values by the nearby unrelated peaks, the convolutional neural networks were applied to extract the important information among features with fully connected layers in each convolution unit (**Figure 1**). In addition, many peak values in the whole spectra data were zero. The introduction of the ReLu function overcame the vanishing gradient problem better than the sigmoid function as suggested for the image classification problems (Krizhevsky et al., 2017) in **Figure 2**. Therefore, the CNN architecture could be an easy and suitable solution for the classification of VRE*fm* and VSEf*m* isolates using whole MS spectra data.

Directly input of whole MS spectra data in the convolutional neural network could be suffered by the exploration of parameters and thus causes the overfitting problem in such networks. It has been demonstrated that dropout techniques are robust to prevent neural networks from overfitting (Hinton et al., 2012; Srivastava et al., 2014). We, therefore, applied the dropout in the later layers of CNN model architecture and yielded significantly better AUROC performance (**Figure 3B**). With a detailed examination of the CNN model architecture, the proposed CNN model in this study remarkably outperformed other ML models, which were developed in our previous works (Wang et al., 2021), for clinical application (**Figure 4**).

Although identification of crucial predictive peaks in VRE*fm* strains may not be essential in clinical application, interpreting CNN models is important to build people's confidence in the system and to facilitate future studies in the exploration of molecular mechanisms behind the resistance. The parameters of deep neural networks are usually large and difficult to be interpreted. However, several techniques such as saliency map (Hong et al., 2015) and Score-CAM (Wang et al., 2020a) have been developed to explain the crucial features in the deep neural networks successfully. In the feature selection section, most of the informative features have been reproducibly identified among 10 CNN models with randomly assigned initial values. That is, our proposed CNN models were robust enough to reach a globally optimal solution to predict VRE*fm* isolates persistently. In this study, many informative features were identified in the ranges from 3,000 to 4,000 Da, which overlapped with those crucial m/z peaks reported in several previous studies (Lasch et al., 2014; Wei et al., 2014; Wang et al., 2021). Moreover, some informative features were present in the VRE*fm* or VSE*fm* isolates, respectively. The findings of VRE*fm*-specific features in 6,602–6,608 Da has been reported that the peak *m/z* 6,603 is specific for vanB-positive VRE*fm* (Griffin et al., 2012). These VRE*fm*-specific features are worthy of further characterization in further investigations.

Indeed, the VRE*fm* prediction model may detect specific resistant clones instead of the resistant mechanisms. Although the investigation into the resistant mechanism of VRE*fm* is a thoughtful question that we cannot totally figure out in the current stage, we conducted strain typing for the VRE*fm* isolates to illustrate the basic molecular composition. We randomly selected 455 VRE*fm* isolates from blood cultures collected in the institutes over time (2002–2015). In the 455 isolates, only 4 isolates (0.88%) were *vanB*, *vanA* was predominant (> 99%). Then, we examined the clones by using a multi-locus sequence for the 455 VRE*fm* isolates. The results showed that a total of 24 ST types were identified (**Supplementary Figure 2**). For testing algorithm on isogenic strains of *E. faecium*, we focused on ST17 as the representative strain and tested ML model performance on discriminating ST17 with *van* gene and ST17 without *van* gene. The ML model attained sensitivity 0.89, specificity 0.8, and AUROC 0.9 in discriminating VSE*fm* ST17 from VRE*fm* ST17. Moreover, the high diverse clone composition for the VRE*fm* isolates has implied that there are some resistance-conferring peptides/proteins in common among various ST types. Given the high diversity of the clones over time, the VRE*fm* prediction model could actually detect the pattern of peaks associated with vancomycin resistance. A total understanding of the underlying mechanism depends on comprehensive identifications of the tens of the informative peaks. Identification of the discriminating peaks would be a tough task that necessaries efforts from the scientific community. In the study, we aimed to develop and validate a novel data preprocessing method that can dig out the implicit information from existing MALDI-TOF spectra for predicting the AST of vancomycin. Meanwhile, the CNN models also provided a list of informative features that are worthy of further molecular investigation to fully understand the underlying mechanism of drug resistance.

There are some limitations to this study. First, the bacterial strains vary according to the environments and locations.

Because our data were collected from two tertiary different medical centers in Taiwan, our trained CNN models may not be universally suitable to predict the VRE*fm* strains in other areas or countries. However, the CNN model algorithm is believed to be a powerful method for classifying VRE*fm* strains in clinical applications. Second, our primary goal was to develop and validate a practical and ready-to-use CNN model in clinical practice using whole MALDI-TOF MS spectra. As mentioned earlier we reported some crucial informative peak features for VRE*fm*, however, it is worthy of further confirmation in the identities for these specific peaks corresponding to which peptide products experimentally.

In conclusion, the CNN model was designed to be used in clinical practice. Based on the design, the input is a whole-cell MALDI-TOF MS spectrum that is routinely used for species identification in the clinical microbiology laboratory. Thus, no additional experiment is needed by clinical microbiologists. Once an isolate is identified as *E. faecium*, the raw MALDI-TOF MS spectrum will be transferred to the CNN model directly. Susceptible or resistant to vancomycin will be predicted in seconds and can be provided to clinical physicians.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Files**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

H-YW, J-TH, J-JL, and J-HH conceived the idea and designed the study. T-TH developed the computational algorithms and performed the data analysis. C-RC and H-CC provided assistance in data analysis and interpretation of the results. H-YW and J-HH interpreted the results and wrote the manuscript. All authors contributed to amending the manuscript and have read the submitted version. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Chang Gung Memorial Hospital (Linkou) [CMRPG3L0401 (J-JL), CMRPG3L0431 (J-JL), and CMRPG3L1011 (H-YW)] and the Ministry of Science and Technology, Taiwan [111-2320-B-182A-002-MY2 (H-YW)].

ACKNOWLEDGMENTS

We thank Dr. Chien-Yu Chen and Dr. Yin-Hung Lin for insightful discussion.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.821233/full#supplementary-material

REFERENCES

Ahmed, M. O., and Baptiste, K. E. (2018). Vancomycin-resistant enterococci: a review of antimicrobial resistance mechanisms and perspectives of human and animal health. *Microb. Drug Resist.* 24, 590–606. doi: 10.1089/mdr.2017.0147

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Chen, T., and Guestrin, C. (2016). "XGBoost: a scalable tree boosting System", in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* KDD'16. (New York, NY, USA: Association for Computing Machinery), 785–794. doi: 10.1145/2939672.2939785

Chung, C.-R., Wang, H.-Y., Lien, F., Tseng, Y.-J., Chen, C.-H., Lee, T.-Y., et al. (2019). Incorporating statistical test and machine intelligence into strain typing of staphylococcus haemolyticus based on matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Front. Microbiol.* 10, 2120. doi: 10.3389/fmicb.2019.02120

CLSI (2020). *Performance Standards for Antimicrobial Susceptibility Testing.* 30th ed. CLSI supplement M100. Wayne, PA: Clinical and Laboratory Standards Institute.

Griffin, P. M., Price, G. R., Schooneveldt, J. M., Schlebusch, S., Tilse, M. H., Urbanski, T., et al. (2012). Use of matrix-assisted laser desorption ionization-time of flight mass spectrometry to identify vancomycin-resistant enterococci and investigate the epidemiology of an outbreak. *J. Clin. Microbiol.* 50, 2918–2931. doi: 10.1128/JCM.01000-12

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. *arXiv:1207.0580 [cs]*. Available online at: http://arxiv.org/abs/1207.0580 (accessed September 15, 2021).

Hong, S., You, T., Kwak, S., and Han, B. (2015). "Online tracking by learning discriminative saliency map with convolutional neural network", in *International Conference on Machine Learning* (PMLR), 597–606. Available online at: https://proceedings.mlr.press/v37/hong15.html (accessed September 15, 2021).

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60, 84–90. doi: 10.1145/3065386

Lasch, P., Fleige, C., Stämmler, M., Layer, F., Nübel, U., Witte, W., et al. (2014). Insufficient discriminatory power of MALDI-TOF mass spectrometry for typing of Enterococcus faecium and Staphylococcus aureus isolates. *J. Microbiol. Method.* 100, 58–69. doi: 10.1016/j.mimet.2014.02.015

Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE.* 86, 2278–2324. doi: 10.1109/5.726791

Mortier, T., Wieme, A. D., Vandamme, P., and Waegeman, W. (2021). Bacterial species identification using MALDI-TOF mass spectrometry and machine learning techniques: a large-scale benchmarking study. *Comput. Struct. Biotechnol. J.* 19, 6157–6168. doi: 10.1016/j.csbj.2021.11.004

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1929–1958. doi: 10.5555/2627435.2670313

Storey, J. D., Bass, A. J., Dabney, A., and Robinson, D. (2020). *qvalue: Q-Value Estimation for False Discovery Rate Control*. doi: 10.18129/B9.bioc.qvalue

Wang, H., Wang, Z., Du, M., Yang, F., Zhang, Z., Ding, S., et al. (2020a). Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks. *arXiv:1910.01279 [cs]*. Available at: http://arxiv.org/abs/1910.01279 (accessed April 13, 2021).

Wang, H.-Y., Chen, C.-H., Lee, T.-Y., Horng, J.-T., Liu, T.-P., Tseng, Y.-J., et al. (2018a). Rapid detection of heterogeneous vancomycin-intermediate staphylococcus aureus based on matrix-assisted laser desorption ionization time-of-flight: using a machine learning approach and unbiased validation. *Front. Microbiol.* 9, 2393. doi: 10.3389/fmicb.2018.02393

Wang, H.-Y., Chung, C.-R., Chen, C.-J., Lu, K.-P., Tseng, Y.-J., Chang, T.-H., et al. (2021). Clinically applicable system for rapidly predicting enterococcus faecium susceptibility to vancomycin. *Microbiol. Spectr.* 9, e0091321. doi: 10.1128/Spectrum.00913-21

Wang, H.-Y., Chung, C.-R., Wang, Z., Li, S., Chu, B.-Y., Horng, J.-T., et al. (2020b). A large-scale investigation and identification of methicillin-resistant Staphylococcus aureus based on peaks binning of matrix-assisted laser desorption ionization-time of flight MS spectra. *Brief. Bioinform.* 22, bbaa138. doi: 10.1093/bib/bbaa138

Wang, H.-Y., Lee, T.-Y., Tseng, Y.-J., Liu, T.-P., Huang, K.-Y., Chang, Y.-T., et al. (2018b). A new scheme for strain typing of methicillin-resistant Staphylococcus aureus on the basis of matrix-assisted laser desorption ionization time-of-flight mass spectrometry by using machine learning approach. *PLoS ONE.* 13, e0194289. doi: 10.1371/journal.pone.0194289

Wang, H.-Y., Li, W.-C., Huang, K.-Y., Chung, C.-R., Horng, J.-T., Hsu, J.-F., et al. (2019). Rapid classification of group B Streptococcus serotypes based on matrix-assisted laser desorption ionization-time of flight mass spectrometry and machine learning techniques. *BMC Bioinform.* 20, 703. doi: 10.1186/s12859-019-3282-7

Wei, W., Li-Jun, W., Wen-Jun, S., Gui, Z., and Xin-Xin, L. (2014). Application of matrix-assisted laser desorption ionization time-of-flight mass spectrometry in the screening of vana-positive enterococcus faecium. *Eur. J. Mass Spectrom.* 20, 461–465. doi: 10.1255/ejms.1298

Weis, C., Cuénod, A., Rieck, B., Dubuis, O., Graf, S., Lang, C., et al. (2022). Direct antimicrobial resistance prediction from clinical MALDI-TOF mass spectra using machine learning. *Nat. Med.* 28, 164–174. doi: 10.1038/s41591-021-01619-9

Wolters, M., Rohde, H., Maier, T., Belmar-Campos, C., Franke, G., Scherpe, S., et al. (2011). MALDI-TOF MS fingerprinting allows for discrimination of major methicillin-resistant Staphylococcus aureus lineages. *Int. J. Med. Microbiol.* 301, 64–68. doi: 10.1016/j.ijmm.2010.06.002

Conflict of Interest: T-TH, H-CC, and J-HH were employed by Taiwan AI Labs.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 20 June 2022
doi: 10.3389/fmicb.2022.784628

Parallel Reaction Monitoring Mass Spectrometry for Rapid and Accurate Identification of β-Lactamases Produced by *Enterobacteriaceae*

*Yun Lu, Xinxin Hu, Jing Pang, Xiukun Wang, Guoqing Li, Congran Li, Xinyi Yang and Xuefu You**

Beijing Key Laboratory of Antimicrobial Agents, Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

The increasing spread of drug-resistant bacterial strains presents great challenges to clinical antibacterial treatment and public health, particularly with regard to β-lactamase-producing *Enterobacteriaceae*. A rapid and accurate detection method that can expedite precise clinical diagnostics and rational administration of antibiotics is urgently needed. Targeted proteomics, a technique involving selected reaction monitoring or multiple reaction monitoring, has been developed for detecting specific peptides. In the present study, a rapid single-colony-processing procedure combined with an improved parallel reaction monitoring (PRM) workflow based on HRAM Orbitrap MS was developed to detect carbapenemases (*Klebsiella pneumoniae* carbapenemase, KPC; imipenemase, IMP; Verona integron-encoded metallo-β-lactamase, VIM; New Delhi metallo-β-lactamase, NDM; and oxacillinase, OXA), extended spectrum β-lactamases (TEM and CTX-M), and AmpC (CMY-2) produced by *Enterobacteriaceae*. Specific peptides were selected and validated, and their coefficients of variation and stability were evaluated. In total, 188 *Enterobacteriaceae* strains were screened using the workflow. Fourteen out of total 19 peptides have 100% specificity; three peptides have specificity >95% and two peptides have specificity ranged from 74∼85%. On the sensitivity, only nine peptides have 95∼100% sensitivity. The other 10 peptides have sensitivity ranged from 27∼94%. Thus, a screening method based on peptide groups was developed for the first time. Taken together, this study described a rapid extraction and detection workflow for widespread β-lactamases, including KPC, IMP, VIM, NDM, OXA, CMY, CTX-M, and TEM, using single colonies of *Enterobacteriaceae* strains. PRM-targeted proteomics was proven to be a promising approach for the detection of drug-resistant enzymes.

Keywords: ***Enterobacteriaceae***, **β-lactamases, specific peptides, detection, PRM**

OPEN ACCESS

Edited by:
Karsten Becker, University Medicine Greifswald, Germany

Reviewed by:
Robert A. Bonomo, United States Department of Veterans Affairs, United States
Miriam Cordovana, Bruker Daltonik GmbH, Germany

***Correspondence:**
Xuefu You
xuefuyou@hotmail.com

Specialty section:
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

Received: *12 January 2022*
Accepted: *03 June 2022*
Published: *20 June 2022*

Citation:
Lu Y, Hu X, Pang J, Wang X, Li G, Li C, Yang X and You X (2022) Parallel Reaction Monitoring Mass Spectrometry for Rapid and Accurate Identification of β-Lactamases Produced by Enterobacteriaceae. Front. Microbiol. 13:784628. doi: 10.3389/fmicb.2022.784628

INTRODUCTION

With the extensive clinical use of carbapenems and β-lactam antibacterial drugs, the incidences of antibiotic resistance have increased. The increasing population of multi-drug resistant (MDR) and extensively drug-resistant strains accompanied with the rapid spread of antibiotic-resistance genes have posed great challenges to clinical anti-bacterial treatment and public health. In the

list of bacteria for which new antibiotics are urgently needed released by the World Health Organization in 2017, carbapenem-resistant *Enterobacteriaceae* (CRE) and extended-spectrum β-lactamase (ESBL)-producing *Enterobacteriaceae*, were included in the Critical group (Priority 1) (WHO, 2017). β-lactamases are a group of bacterial enzymes that can inactivate β-lactam antibiotics, resulting in the loss of antibacterial activity (Eliopoulos and Karen, 2001). β-lactamases are currently divided into four classes: A, B, C, and D according to the Ambler classification, based on their primary structure (Bush and Jacoby, 2010). Carbapenems are generally regarded as the last treatment choice for serious bacterial infections. Carbapenemases are β-lactamases with versatile hydrolytic capacities: the A and D class carbapenemases are serine-type hydrolases, such as *Klebsiella pneumoniae* carbapenemase (KPC), and oxacillinase (OXA). The B class carbapenemases are metallo-hydrolases, such as New Delhi metallo-β-lactamase (NDM), imipenemase (IMP), and Verona integron-encoded metallo-β-lactamase (VIM). ESBLs belong to the class A and D β-lactamases. ESBLs such as the TEM (ampicillin resistance) and CTX-M (cefotaxime resistance) groups belong to class A β-lactamases (Pitout et al., 2005). AmpC β-lactamases such as CMY-2 belong to class C of carbapenemases (Philippon et al., 2002). *Enterobacteriaceae*, including *Escherichia coli*, *K. pneumoniae*, and *Enterobacter cloacae,* which carry several types of β-lactamases, represent a great challenge to the clinical diagnostics and treatment of various infections (Duin, 2017).

Early diagnosis and effective drug treatment are key strategies to deal with the antibiotic-resistance problem (Rodríguez-Baño et al., 2018). Thus, rapid and accurate detection methods are urgently needed in clinical practice. A short testing time and accurate diagnosis will assist in providing an appropriate antibiotic treatment in time. For the last few decades, traditional approaches for the detection of β-lactamases, such as standard disk-diffusion procedure, broth microdilution, and agar dilution, have been used, which are time-consuming and can only determine the drug-resistance property but not the β-lactamase type. Synergy testing is accurate and allow to identify the carbapenemase classes, which is the info needed for therapeutic choices. Polymerase chain reaction (PCR) are very sensitive and fast and can be used also from primary samples (Dallenne et al., 2010). Whole genome sequencing, which can accurately identify the genotypes of β-lactamases, is still not practical for usual clinical application because of the high price. However, since drug-resistant enzymes are the products of regulated expression, the detection of a gene may not be correlated with the successful expression of the β-lactamases. Carba NP test based on the detection of enzyme activity (Vasoo et al., 2013) are rapid but not specific for a single type of carbapenemase. The mCIM and eCIM, phenotypic detection methods based on carbapenem inactivation methods, can detect carbapenemases in *Enterobacteriaceae* and *Pseudomonas aeruginosa* and differentiate metallo-beta-lactamases from serine carbapenemases in *Enterobacteriaceae* (Tsai et al., 2020).

The direct detection and quantification of β-lactamases have become easier with the development of protein detection and quantitation techniques in recent years. Lateral flow immunoassay methods based on antigen-antibody reaction (Boutal et al., 2017) are limited to the types of antibodies, but actually they cover the vast majority of carbapenemases found in clinical microbiology routine, beyond being extremely fast. For the past few decades, liquid chromatography with tandem mass spectrometry (LC-MS/MS) has been widely used in various fields of protein analysis, biochemical analysis, natural product analysis, and drug and food analysis among other areas (Suh et al., 2017). LC-MS/MS has gradually become one of the most popular analytical tools for protein detection. Shotgun proteomics has been used to identify wild-type and resistant strains of the pathogen *Acinetobacter baumannii* (Chang et al., 2013). Additionally, a capillary electrophoresis-electrospray ionization-tandem mass spectrometry bottom-up proteomics workflow has been established for the identification of OXA-48 and KPC (Fleurbaaij et al., 2014). MALDI-TOF MS has been used to identify for carbepanemase detection with different analytical approaches (Hleba et al., 2021). Additionally, the bottom-up proteomics approach has also been applied to identify CTX-M ESBLs (Fleurbaaij et al., 2017). Recently, targeted LC-MS/MS based on selected reaction monitoring (SRM) and multiple reaction monitoring (MRM) using triple quadrupole mass analyzer, and parallel reaction monitoring (PRM) techniques using on high resolution/accurate mass has been used in β-lactamase testing. Specific peptides of *A. baumannii* identified *via* LC-MS/MS profiling have been used to classify clinical isolates (Honghui et al., 2016). Targeted high-resolution MS assays have been developed for the detection of KPC, OXA-48, NDM, and VIM enzymes (Wang et al., 2017, 2019; Foudraine et al., 2019; Strich et al., 2019). However, there are still many β-lactamases that have not been detected by targeted proteomics. Traditional SRM-MS on a triple quadrupole mass spectrometer is limited in the complex sample analysis due to the mass filtering and low resolution quadrupole, and more method development time is needed to define transitions (precursor/product ion pairs). In our study, a system comprising a rapid sample-processing procedure combined with improved PRM using HRAM (High resolution accurate mass) Orbitrap MS was developed to detect carbapenemases (KPC, IMP, VIM, NDM, and OXA), ESBLs (TEM and CTX-M), and AmpC (CMY-2) using a single colony of *Enterobacteriaceae* strains.

MATERIALS AND METHODS

Strains and Culture Conditions

A total of identified and subcultured 192 *Enterobacteriaceae* strains were used in this study (**Supplementary Table 3**), including *K. pneumoniae* American Type Culture Collection (ATCC) BAA-2146 and 191 clinically isolated strains. The selected strains comprised 73 *E. coli*, 83 *K. pneumoniae,* 25 *E. cloacae* strains, five *Klebsiella oxytoca*, and six *Citrobacter freundii*. Four of these isolates were used for the test development process, and 188 were used for the test validation phase. All the strains were cultured in Luria-Bertani agar plates at 37°C overnight. All the strains were stored in the Collection Center

of Pathogen Microorganism of Chinese Academy of Medical Sciences in China.

Peptide Preparation

Single colonies (diameter >2 mm) were picked using a micropipette tip or a 10 μL loop, and resuspended in 200 μL of 50 mM ammonium bicarbonate (MS grade; Merck, Germany), sonicated for 1 min (3 s of sonication, 6 s of rest), centrifuged at 12,000 × *g* for 2 min and heated at 95°C for 5 min, after which the buffer was removed using 10K Nanosep centrifugal device with Omega membrane (Pall Corporation, Port Washington, NY, United States). Ammonium bicarbonate buffer (50 mM) was added along with sequencing grade trypsin (Promega Corporation, Madison, WI, United States), and the solution was microwaved in a water bath followed by heat treatment at 55°C. The peptide concentration was measured using the Pierce™ Quantitative Colorimetric Peptide Assay kit (Thermo Fisher Scientific, Waltham, MA, United States).

Nano LC-MS/MS

Data-dependent analysis was performed on the Thermo Scientific Orbitrap Fusion Lumos platform coupled with an EASY-nLC 1200 system (Thermo Fisher Scientific, San Jose, CA, United States) to build a spectral library. The digests were separated by the trap column [ReproSil-Pur 120 C18-AQ (3 μm, Dr. Maisch GmbH, Ammerbuch, Germany); 20 × 0.05 mm] followed by a C18 column [ReproSil-Pur 120 C18 (1.9 μm, Dr. Maisch GmbH, Ammerbuch, Germany); 120 × 0.15 mm] at a flow rate of 600 μL/min. The solvent buffer A comprised water with 0.1% formic acid, and solvent B comprised 80% acetonitrile with 0.1% formic acid. After sample loading, the gradient was initiated with 11% of buffer B, and then increased from 11 to 13% of buffer B for 2 min. The gradient increased up to 32% of buffer B in 16 min, and then to 42% in 7 min. Finally, the gradient was increased to 95% of buffer B in 1 min and was maintained for 4 min. The MS parameters were as follows: MS1 (Orbitrap analysis; mass range, approximately 350–1,550 m/z; resolution, 120,000; AGC target, 5×10^5; RF lens, 50%; maximum injection time, 50 ms), MS2 [high-energy collisional dissociation (HCD) collision energy, 32%; maximum injection time, 22 ms; AGC, 5×10^4; isolation window, 1.6 Da; Orbitrap solution, 15,000]. For database analyses, these raw data were searched against a combined bacterial antimicrobial resistance database downloaded from the National Center for Biotechnology Information (NCBI; National Institutes of Health, Bethesda, MD, United States) using the Thermo Scientific™ Proteome Discoverer™ version 2.2 (PD2.2) software. Trypsin was specified for protein digestion with two missed cleavages allowed for each peptide. The search parameters were set as previously described (Espadas et al., 2017).

Peptide Selection

The protein sequences of KPC, VIM, IMP, NDM, OXA, TEM, CMY, and CTX-M were downloaded from the NCBI database[1] to build an amino acid sequence library. Tryptic peptides were searched and aligned using the protein basic local alignment search tool (BLAST[2]) to ensure uniqueness. The peptides with missed-cleaved, two neighboring basic amino acids at either cleavage site (KK, RR, KR, and RK), and W and M residues were excluded.

Targeted Proteomics and Data Analysis

Targeted proteomics was performed on the Thermo Scientific Orbitrap Fusion Lumos platform coupled with an EASY-nLC 1200 platform. QQDLVDYSPVSEK and VDAGDEQLER served as internal reference. The columns used and elution gradient were the same as mentioned above. PRM parameters: MS1 spectrum (Orbitrap analysis; resolution, 60,000; mass range, approximately 350–2000 m/z; RF lens, 30%; AGC target, 2.0×10^5; maximum injection time, 50 ms) and MS2 analysis (HCD; collision energy, 30%; AGC, 5.0×10^4; maximum injection time, 54 ms for specific peptides or 22 ms for synthetic isotope labeled (SIL) peptides; Orbitrap resolution, 30,000 for specific peptides or 7,500 for SIL; isolation window, 1.4 Da). Data on the peptides [retention time (RT), m/z, and charge] were imported into the MS method. For data analysis, the acquired data were analyzed using Skyline 20.1.0.155 (MacCoss Lab Software, University of Washington, Seattle, WA, United States) (MacLean et al., 2010). The amino acid sequences of the drug-resistant enzymes downloaded from the NCBI were imported as the background library. Data-dependent acquisition (DDA) raw data were imported to build the spectra database. After the amino acid sequences of specific and SIL were inserted, the PRM data were imported and for each targeted peptide, the ratio between the peak area of the endogenous peptide and that of the SIL was calculated, and the relative concentration of targeted peptides was calculated based on the SIL with fixed quantity. Whether the CRE/ESBL enzymes were defined as positive or negative depended on the peptide when the following criteria were met: an RT similar to that of the SIL, library dot product (dotp) > 0.8 and ratio dot products (rdotp) > 0.95. Coefficients of variation (CV) were calculated, and stability of the peptides was evaluated by performing three freeze-thaw cycles of the peptides and storage of the peptides for 0, 1, 3, and 4 days in the sample holder (10°C).

Detection of Drug-Resistant Genotypes *via* Polymerase Chain Reaction

Multi-drug resistant genes were analyzed *via* PCR (Dallenne et al., 2010) using the GoTaq Green Master Mix (Promega). The primers and parameters used are listed in **Supplementary Table 1**.

RESULTS

Selection of Unique Peptides

Amino acid sequences of the enzymes were downloaded from the NCBI,[3] and potential peptides were evaluated using

[1]https://www.ncbi.nlm.nih.gov/pathogens/refgene/

[2]https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome

[3]https://www.ncbi.nlm.nih.gov/protein/

PeptideCutter.[4] In view of the varying responses of peptides analyzed *via* MS, four strains (**Table 1**) were used to evaluate the ionization capabilities of peptides *via* DDA, and the specificity of the peptides was assessed by performing BLASTp searches. Unique peptides with high signal stability, appropriate RT, and relatively stable amino acid residues were chosen as peptides markers for KPC, IMP, VIM, NDM, OXA, CMY, CTX-M, and TEM. Val ($13C_5$, 15N), Gly (15N), and Ala ($13C_3$, 15N) were used to label the peptides (**Table 2**). Other candidate peptides detected are listed in **Supplementary Table 2**. The data are deposited in the PRIDE repository, accession number PXD028791.

Parallel Reaction Monitoring Assay Development

To develop a rapid method for peptide detection, a 30-min Nano LC-MS/MS method was developed, with parameters for peptide markers as shown in **Table 2**. Different Orbitrap resolutions for specific and SIL were used to improve identification speed, quantity and quality (Stopfer et al., 2021). **Figure 1** shows the workflow of the rapid detection method for β-lactamases. The raw data were analyzed using Skyline, and the library dotp, rdotp, and R ratio values were exported. According to the library dotp, rdotp, and R ratio values obtained from the strains, the following rules were set: library dotp > 0.8, rdotp > 0.95, and RT similar to that of SIL (**Figure 2**). A wash procedure for 30 min was performed after each sample to avoid false-positive results caused by the carryover effect.

Optimization of Rapid Digestion Conditions

To obtain high-quality spectra while reducing the digestion time, protein solutions with trypsin were microwaved for 5, 10, and 15 min, as well as heated in a water bath at 55°C for 15, 30, and 45 min after being microwaved for 2 min separately. Peptides were collected and analyzed *via* Nano LC-MS/MS. The total processing time was <1 h. As shown in **Supplementary Figure 1**, the peptides could be detected for all digestion conditions even after being microwaved for 5 min. For most peptides, an increase in digestion time led to an increase in abundance; however, the digestion conditions had no effect on the peptides LAEAEGNEIPTHSLEGLSSSGDAVR, VQATNSFSGVNYWLVK, IINHNLPVK, and NSFGGVNYW LVK. To ensure that all peptides were identified under optimal conditions, microwave treatment for 2 min combined with heat treatment using a 55°C water bath for 30 min was performed in subsequent experiments.

[4]https://web.expasy.org/peptide_cutter/

Reproducibility and Stability Tests

The reproducibility of applications of the peptide markers was evaluated using four strains that were positive for KPC, IMP, VIM, NDM, OXA, CMY, CTX-M, and TEM based on the CV of six replicates in 1 day and in 3 different days (**Table 3**). The CVs of endogenous contents were determined using SIL (**Table 2**). The CVs of 16 peptides were <30%. While the CV of IINHNLPVK wasn't acquired as signal miss in some of the samples. No carry-over effects were observed for any of the selected peptides.

Stability is an important property that should be evaluated in method development. The SIL were added to the peptide solutions of *E. coli* DH5α that did not contain β-lactamases. The peptide stability after three freeze-thaw cycles and stability in the sample holder were measured separately. Six replicates were performed for all experiments. As shown in **Supplementary Figure 2**, the contents of DGDELLLIDTAWGAK and VGYIELDLNSGK decreased significantly after three freeze-thaw cycles (<80%), whereas the contents of LIAQLGGPGGVTAFAR and APLILVTYFTQPQPK decreased slightly (approximately 80–90%). These results suggest that repeated freezing and thawing of these four peptides should be avoided during use. Regarding stability in the sample holder for 0, 1, 3, and 4 days, the content of DGDELLLIDTAWGAK decreased significantly (70%) in 1 day, and down to 6% on the fourth day (**Supplementary Figure 3**). The content of VGYIELDLNSGK decreased significantly (64%) on the third day, whereas the content of APLILVTYFTQPQPK reduced to approximately 82% on the fourth day. The results suggest that these peptides should be held for short durations in the sample holder.

Method Validation

After the preliminary evaluation of the rapid detection method, blinded testing of 188 clinical strains was conducted. Strains of *E. coli*, *K. pneumoniae, E. cloacae* strains, *K. oxytoca*, and *C. freundii* were included in the assays. As shown in **Table 4**, all the β-lactamases tested were detected both *via* LC-MS/MS and PCR. Most of the peptide markers for the β-lactamases showed 100% specificity except for LDEGVYVHTSFEEVNGWGVVPK(IMP, 85%), TEPTLNTAIPGDPR(CTX-M-1 group and 9 group, 97%), APLIL VTYFTQPQPK(CTX-M-1 group, 99%), SDLVNYNPIAEK

TABLE 1 | β-lactamase information of isolates used in method development.

Name	KPC	NDM	VIM	IMP	OXA	TEM	CMY	CTX-M
Klebsiella pneumoniae ATCC BAA-2146 (Kwon et al., 2016)	−	+	−	−	+	+	+	+
Klebsiella pneumoniae 1705[a]	+	+	−	−	+	+	−	+
Klebsiella pneumoniae 17-R66[a]	−	−	−	+	−	+	−	+
Klebsiella pneumoniae 17-R42[a]	−	−	+	+	−	+	−	−

[a]*Measured by PCR and DNA sequencing.*

KPC, Klebsiella pneumoniae carbapenemase; IMP, imipenemase; VIM, Verona integron-encoded metallo-β-lactamase; NDM, New Delhi metallo-β-lactamase; OXA, oxacillinase; CTX-M, β-lactamase against cefotaxime.

TABLE 2 | Selected target peptides for the rapid detection of multi-drug resistance enzymes.

Peptide	Enzyme	Genotype[a]	Labeled site	Charge	m/z (unlabeled)	RT (min)
AAVPADWAVGDK	KPC	1∼95	GLY(15N)	2	600.3064	15.4
SQQQAGLLDTPIR	KPC	1∼95 except for 13, 45, 59	GLY(15N)	2	713.8861	16.7
LVVPSHSEVGDASLLK	IMP-1	1, 5, 7, 10, 28, 30, 34, 40, 42, 43, 52, 55, 60, 61, 66, 70, 73, 76, 77, 79, 81, 85, 88	GLY(15N)	2	825.9567	16.1
VQATNSFSGVNYWLVK	IMP-1	1, 3, 6, 10, 25, 30, 34, 40, 42, 52, 55, 60, 61, 66, 70, 76∼80, 88	GLY(15N)	2	906.9676	22.9
NSFGGVNYWLVK	IMP-4	4, 26, 38, 59, 89	GLY(15N)	2	692.3564	22.6
LDEGVYVHTSFEEVNGWGVVPK	IMP	1, 3, 4, 5, 6, 7, 15, 25, 28, 29, 34, 38, 51, 52, 60, 61, 62, 59, 70, 79, 81, 82, 85	GLY(15N)	3	821.0727	22.2
LAEAEGNEIPTHSLEGLSSSGDAVR	VIM	1, 4, 5, etc.	GLY(15N)	3	847.0805	16.33
DGDELLLIDTAWGAK	VIM	All except for 7, 13, 47, 61, 69	GLY(15N)	2	808.9120	26.71
AFGAAFPK	NDM	1∼31	GLY(15N)	2	404.7212	14.6
NNGLTEAWLESSLK	OXA-1 family	1, 4, 31, 47, 224, 320, 392, 534, 675(oxa-1 family)	GLY(15N)	2	781.3965	22.3
IINHNLPVK	OXA-1 family	1, 4, 31, 47, 224, 320, 392, 534, 675(oxa-1 family)	VAL (13C5, 15N)	3	349.8818	8.8
ADIANNHPVTQQTLFELGSVSK	CMY-2 family	2, 4, 5, 6, 7, etc.	GLY(15N)	2	1185.108	19.5
TLQQGIALAQSR	CMY-2 family	2, 4, 5, 6, 7, etc.	GLY(15N)	3	429.2454	14.2
QLTLGHALGETQR	CTX-M-9 group	9, 13, 14, 17, 21, 19, 24, 65, 81, etc.	GLY(15N)	3	475.2599	13.3
TEPTLNTAIPGDPR	CTX-M	All ctx-m genotypes except for 4, 6, 7, 19, 23, 35, 42, 52, 54, 58, 62, 74, 87, 93, 99, 117, 126, 144, 147, 151, 155, 157, 168, 204, 219, 212, 221	GLY(15N)	2	741.3834	15.1
LIAQLGGPGGVTAFAR	CTX-M-9 group	9, 13, 14, 16, 21, 17, 19, 81, etc.	GLY(15N)	2	764.4357	20.8
APLILVTYFTQPQPK	CTX-M-1 group	1, 3, 10, 12, 15, etc.	Ala (13C3,15N)	2	858.4902	26.8
SDLVNYNPIAEK	CTX-M-1 group	1, 2, 5, 12, 15, etc.	Ala (13C3,15N)	2	681.8486	16.1
VGYIELDLNSGK	TEM	All except for 60, 139, 178, 210	GLY(15N)	2	654.3457	19.56

[a]*All the genotype-matching results are based on data obtained from the NCBI ANTIMICROBIAL RESISTANCE GENE database. With an increase in data, the results will change. m/z, mass to charge ratio; RT, retention time.*

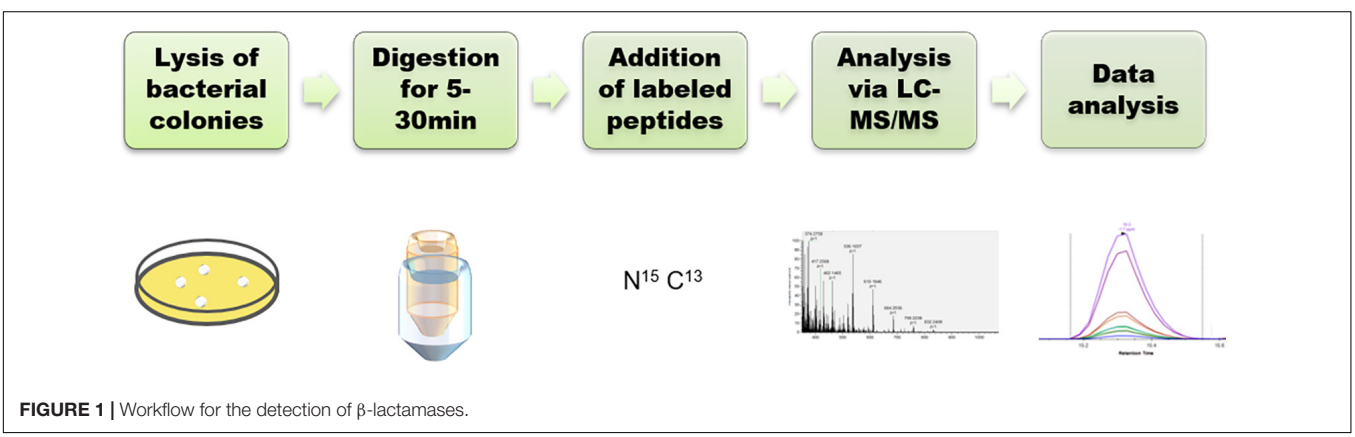


FIGURE 1 | Workflow for the detection of β-lactamases.

(CTX-M-1 group, 95%), and VGYIELDLNSGK(TEM, 74%). Peptide markers for KPC(AAVPADWAVGDK), IMP-1, VIM-1, and NDM showed 100% sensitivity. However, the positive sensitivities for SQQQAGL LDTPIR(KPC), NSFGGVNYWLVK(IMP-4), LDEGVYVHTS FEEVNGWGVVPK(IMP), NNGLTEAWLESSLK(OXA-1), II NHNLPVK(OXA-1), ADIANNHPVTQQTLFELGSVSK(CMY-2), TLQQGIALAQSR(CMY-2), QLTLGHALGETQR(CTX-M-9

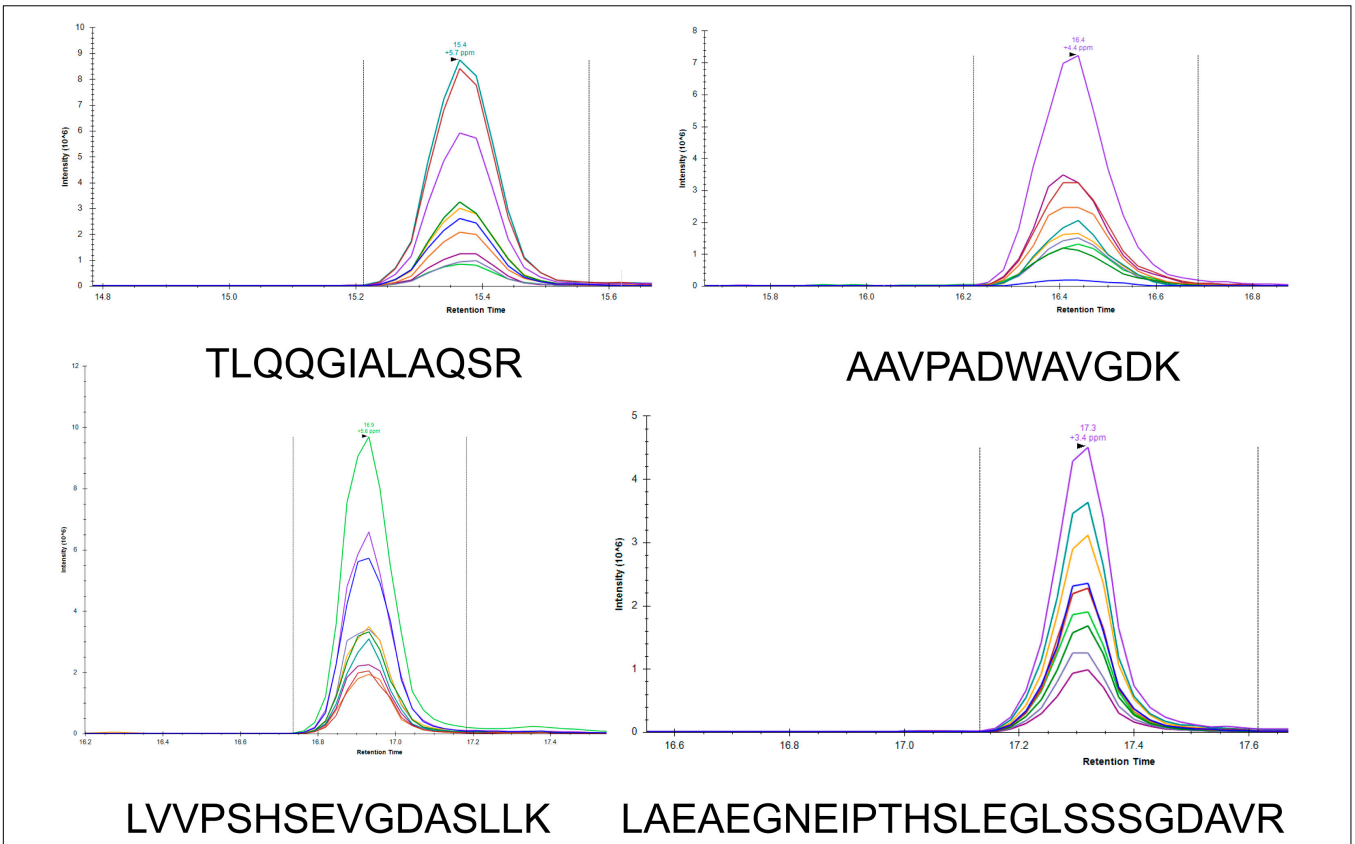


FIGURE 2 | LC-MS/MS chromatograms of the peptides (TLQQGIALAQSR, AAVPADWAVGDK, LVVPSHSEVGDASLLK, and LAEAEGNEIPTHSLEGLSSSGDAVR) in Skyline. Different colors represent different fragment ions.

group), LIAQLGGPGGVTAFAR(CTX-M-9 group), TEPTLN TAIPGDPR(CTX-M-9 group, CTX-M-1 group partial), APLILVTYFTQPQPK(CTX-M-1 group), SDLVNYNPIAEK (CTX-M-1 group partial), and VGYIELDLNSGK(TEM) were 98% (49/50), 58% (14/24), 92% (24/26), 78% (14/18), 78% (14/18), 27% (6/22), 82% (18/22), 39% (27/70), 81% (57/70), 98% (116/118), 98% (57/58), 76% (44/58), and 74% (67/90), respectively. Peptide values were calculated using the labeled peptides that are listed in **Table 4**.

DISCUSSION

For the past few decades, β-lactam antibiotics have been one of the first-choice drugs for the treatment of several serious infectious diseases caused by Gram-negative bacteria (Bush and Bradford, 2016). However, the increasing rate of drug-resistance in bacteria has far exceeded the rate of development of new antibiotics at present. The spread of β-lactamases is mainly attributed to the presence of β-lactamase genes on plasmids, and the unreasonable use of antibiotics. Therefore, a timely detection of the type of β-lactamases will promote appropriate clinical antibiotic choice and inhibit the spread of drug-resistant genes (Iovleva and Doi, 2017). With the development of HRAM Orbitrap MS and supporting quantitative methods, the application of detection approaches based on specific peptides has gradually received attention. Initially, shotgun proteomics was used for β-lactamase detection and the proteins were detected in a single run; however, this depends on the database and results in a poor accuracy (Fleurbaaij et al., 2017). Subsequently, quantitative proteomics methods such as SRM and MRM have been applied to β-lactamases analysis, and detection based on specific peptides has proved to be feasible (Wang et al., 2017). Using LC-MS/MS to detect β-lactamases is a more direct approach than using PCR or disk-diffusion methods, and it is less hindered by multiple problems such as false-positive results. However, for peptide detection, the preparation procedure of peptides must be optimized to reduce the time taken.

In our study, we developed a rapid preparation procedure based on a single colony, thereby omitting the amplification process. Isolating bacteria using agar plates is the first step for isolating organisms from all types of clinical samples. As long as colonies are acquired, the identification process may be initiated. The procedures involving the reduction and alkylation of sulfhydryl groups were also removed as they do not significantly affect the digestion of targeted peptides. Moreover, a simplified Filter-aided sample preparation (FASP) method was used to ensure digestion efficiency. Previous studies have confirmed that microwave treatment for short periods can effectively digest the targeted peptides (Strich et al., 2019). In our study, we tested

TABLE 3 | Intensity ratios and coefficients of variation of peptides.

Enzyme	Peptide	CV (%) in 1 day	CV (%) in 3 different day
KPC	AAVPADWAVGDK	23	25
KPC	SQQQAGLLDTPIR	12	12
IMP-1	LVVPSHSEVGDASLLK	8	11
IMP-1	VQATNSFSGVNYWLVK	10	12
IMP-4	NSFGGVNYWLVK	9	20
IMP	LDEGVYVHTSFEEVNGWGVVPK	9/14	9/20
VIM	LAEAEGNEIPTHSLEGLSSSGDAVR	5	8
VIM	DGDELLLIDTAWGAK	1	2
NDM	AFGAAFPK	14/27	23/10
OXA-1 family	NNGLTEAWLESSLK	6/21	5/17
OXA-1 family	IINHNLPVK	*	*
CMY-2 family	ADIANNHPVTQQTLFELGSVSK	15	17
CMY-2 family	TLQQGIALAQSR	13	14
CTX-M-9 group	QLTLGHALGETQR	11	15
CTX-M	TEPTLNTAIPGDPR	6/6/17	9/13/9
CTX-M-9 group	LIAQLGGPGGVTAFAR	15/16	14/24
CTX-M-1 group	APLILVTYFTQPQPK	37/72	33/57
CTX-M-1 group	SDLVNYNPIAEK	36/42	32/36
TEM	VGYIELDLNSGK	3/6/13	11/5/4

*CV, coefficient of variation; *IINHNLPVK was not detected in both samples.*

different digestion conditions for our selected peptides. Results showed that all the targeted peptides could be detected even after microwave treatment for only 5 min, which can greatly reduce the digestion time. For the LC-MS/MS procedure, a 30 min LC method combined with a PRM targeted method was used to identify the peptides. PRM can enable identification of multiple peptides in a high resolution and high mass accuracy mode (Navin, 2015). In contrast with a previous PRM detection method, we used a lower resolution for labeled peptides and higher resolution for targeted peptides to reduce the scanning time and improve the MS/MS quality as the concentration of SIL was high (Stopfer et al., 2021). By optimizing the detection methods, we have obtained a series of peptides with varying properties. Overall, a 30∼60 min preparation procedure, a 30 min LC-MS/MS procedure and a 10 min data processing were determined for the detection of β-lactamases.

For KPC, AAVPADWAVGDK and SQQQAGLLDTPIR were selected as peptides markers (**Table 2** and **Supplementary Table 2**); LALEGLGVNGQ, LTLGSALAAPQR, and APIVLAVYTR were previously analyzed using Agilent 6540 Q-TOF (Wang et al., 2017). NALVPWSPISEK was detected for the first time but was not selected as a marker owing to its low dotp values. Both of the sensitivity and specificity of AAVPADWAVGDK was 100% as it exists in all the genotypes of bla_{KPC}, indicating the possibility of becoming a peptide marker for KPC. But for SQQQAGLLDTPIR, the sensitivity was lower as the absence of bla_{KPC-13}, bla_{KPC-45}, and bla_{KPC-59}. This study shows the first successful detection of IMP *via* LC-MS/MS. The results proved that LVVPSHSEVGDASLLK, VQATNSFSGVNYWLVK, and NSFGGVNYWLVK could effectively be used to identify IMP and distinguish the subtypes of bla_{IMP-1} and bla_{IMP-4}. LDEGVYVHTSFEEVNGWGVVPK showed lower specificity for IMP detection. And for NSFGGVNYWLVK, it is inexplicable that the discrepant results were verified to be bla_{IMP-4} positive,

TABLE 4 | Results of parallel reaction monitoring compared to those of polymerase chain reaction (PCR) in a validation set.

Enzyme	Peptide	Sensitivity%, CI%, (n MS positive/n PCR positive)	Specificity%, CI%, (n MS negative/n PCR negative)	Sensitivity%, CI%, (n MS positive/n PCR positive) group	Specificity%, CI%, (n MS negative/n PCR negative) group
KPC	AAVPADWAVGDK	100, 91–100 (50/50)	100, 97–100 (138/138)	100, 91–100 (50/50)	100, 97–100 (138/138)
KPC	SQQQAGLLDTPIR	98, 88–100 (49/50)	100, 97–100 (138/138)		
IMP-1	LVVPSHSEVGDASLLK	100, 60–100 (8/8)	100, 97–100 (180/180)	100, 60–100 (8/8)	100, 97–100 (180/180)
IMP-1	VQATNSFSGVNYWLVK	100, 60–100 (8/8)	100, 97–100 (180/180)		
IMP-4	NSFGGVNYWLVK	58, 37–77 (14/24)	100, 97–100 (164/164)	58, 37–77 (14/24)	100, 97–100 (164/164)
IMP	LDEGVYVHTSFEEVNGWGVVPK	92, 73–99 (24/26)	85, 79–90 (138/162)	92, 73–99 (24/26)	85, 79–90 (138/162)
VIM-1	LAEAEGNEIPTHSLEGLSSSGDAVR	100, 20–100 (2/2)	100, 97–100 (186/186)	100, 20–100 (2/2)	100, 97–100 (186/186)
VIM-1	DGDELLLIDTAWGAK	100, 20–100 (2/2)	100, 97–100 (186/186)		
NDM	AFGAAFPK	100, 80–100 (20/20)	100, 97–100 (168/168)	100, 80–100 (20/20)	100, 97–100 (168/168)
OXA-1	NNGLTEAWLESSLK	78, 52–93 (14/18)	100, 97–100 (170/170)	94, 71–100 (17/18)	100, 97–100 (170/170)
OXA-1	IINHNLPVK	78, 52–93 (14/18)	100, 97–100 (170/170)		
CMY-2	ADIANNHPVTQQTLFELGSVSK	27, 12–50 (6/22)	100, 97–100 (166/166)	86, 64–96 (19/22)	100, 97–100 (166/166)
CMY-2	TLQQGIALAQSR	82, 59–94 (18/22)	100, 97–100 (166/166)		
CTX-M-9 group	QLTLGHALGETQR	39, 27–51 (27/70)	100, 96–100 (118/118)	81, 70–89 (57/70)	100, 96–100 (118/118)
CTX-M-9 group	LIAQLGGPGGVTAFAR	81, 70–89 (57/70)	100, 96–100 (118/118)		
CTX-M	TEPTLNTAIPGDPR	98, 93–100 (116/118)	97, 89–100 (68/70)	98, 93–100 (116/118)	97, 89–100 (68/70)
CTX-M-1 group	APLILVTYFTQPQPK	98, 89–100 (57/58)	99, 95–100 (129/130)	98, 90–100 (57/58)	94, 88–97 (122/130)
CTX-M-1 group partial	SDLVNYNPIAEK	76, 83–86 (44/58)	95, 89–98 (123/130)		
TEM	VGYIELDLNSGK	91, 83–96 (89/98)	74, 64–83 (67/90)	91, 83–96 (89/98)	74, 64–83 (67/90)

PCR, polymerase chain reaction; MS, mass spectrometry.

while the Skyline map fragments irons of the discrepant results was different from the positive ones. This is the first time this phenomenon has been observed, and it may be related to the fragmentation of the peptide. Notably, LVVPSHSEAGDASLLK (IMP-4), whose A was replaced by V in IMP-1 (Δ14Da), was also detected and could be separated *via* high-resolution MS. LAEAEGNEIPTHSLEGLSSSGDAVR and DGDELLLIDTAWGAK were used to detect VIM-1, and both of them showed a better subtype coverage. However, as the lack of bla_{VIM-1} strains collected, only 2 bla_{VIM-1} positive strains were identified in 188 strains. With the continuous collection of bla_{VIM-1} positive strains, the effectiveness of detection by peptide markers can be better verified in the future. AFGAAFPK was selected as the marker as it exists in all NDM subtypes currently listed by the NCBI, and was also detected *via* the MRM targeted method (Wang et al., 2019). bla_{OXA} is widely distributed in *Enterobacteriaceae*. Regarding the bla_{OXA-1} family, four peptides were detected, and NNGLTEAWLESSLK and IINHNLPVK, which were specific for the bla_{OXA-1} family were used for screening. However, according to the reproducibility results, the LC-MS/MS signal of IINHNLPVK was not stable in the detection process. bla_{CMY} is an AmpC type ESBL gene. ADIANNHPVTQQTLFELGSVSK and TLQQGIALAQSR showed 27 and 82% sensitivity for the strains used in this study. However, for TEM (including TEM-1, TEM-2, and variants listed on the NCBI ANTIMICROBIAL RESISTANCE GENE database), LLTGELLTLASR, SALPAGWFIADK, IHYSQNDLVEYSPVTEK, QIAEIGASLIK, VDAGQEQLGR, VDAGQEQLGRR, and VGYIELDLNSGK were identified in DDA experiments (**Supplementary Table 2**). QIAEIGASLIK and VGYIELDLNSGK were choose as peptides markers, and QIAEIGASLIK was removed for its poor spectrum. For the 188 strains detected, VGYIELDLNSGK showed 91% sensitivity with lower specificity as 74% (67/90). Overall, this is the first study to use a PRM-based LC-MS/MS system to detect OXA, CMY, and TEM. $bla_{\mathrm{CTX-M}}$ ESBL is a large group of ESBLs with an increasing number of subtypes. According to genetic structure, $bla_{\mathrm{CTX-M}}$ enzymes are divided into four groups (CTX-M-1, CTX-M-2, CTX-M-9, and CTX-M-25) (Canton et al., 2012). Among them, the CTX-M-15 (CTX-M-1 group) and CTX-M-14 (CTX-M-9 group) are by far the most prevalent enzymes. In our study, only APLILVTYFTQPQPK showed 98% sensitivity for the CTX-M-1 group. QLTLGHALGETQR and LIAQLGGPGGVTAFAR only existed in certain strains of the CTX-M-9 group, whereas SDLVNYNPIAEK showed 76% sensitivity for the CTX-M-1 group. Additionally, TEPTLNTAIPGDPR could identify 98% of the strains containing CTX-M-1 or CTX-M-9 groups. The reason for the relatively lower sensitivity may be that the peptides are not the representative peptides for the entire genotype group, but only for the partial group due to the complexity of the variants. It may not be possible to use single peptide to achieve 100% sensitivity. As the overall detection sensitivity for each drug resistant enzyme is more important than the detection sensitivity of individual peptide, to overcome this complexity of huge number of variants, sensitivity and specificity based on peptides groups were calculated as shown in **Table 4**. The sensitivity of KPC, OXA, CMY-2, CTX-M-9 group, and CTX-M-1 peptides groups raised to 100, 94, 76, 81, and 98%. In addition, much more peptide markers need to be effectively detected and added for the group detection in the future study. In addition, several variants were observed for the peptides of $bla_{\mathrm{CTX-M}}$, such as LGVALIDTADNTQVLYR, LGVALINTADNTQTLYR, and LGVALINTADNSQILYR, which are similar in terms of m/z and RT, representing great challenges to detection *via* MS. Therefore, to improve the sensitivity of detection *via* LC-MS/MS, an overall analysis is suggested using all the peptides in one method. Further studies are required to identify specific peptides for one group or the entire $bla_{\mathrm{CTX-M}}$ group.

The dotp is a measure of similarity between the fragment ratio of the endogenous and library peptide. The rdotp is a measure of similarity between the fragment ratio of the endogenous and SIL peptide. The criterium for judging positivity by a previous study was rdotp value > 0.95 (Foudraine et al., 2019). When we re-analyzed the results of the targeted proteomics experiment, we found that unified screening rules might not have been suitable for all peptides. For many peptides such as LDEGVYVHTSFEEVNGWGVVPK (IMP), DGDELLLIDTAWGAK (VIM-1), and TEPTLNTAIPGDPR (CTX-M-1 group partial), a rdotp value > 0.95 is not enough to screen for positive strains; therefore, library dotp values > 0.8 are required. For QLTLGHALGETQR (CTX-M-9), 100% (70/70) strains were positive for library dotp > 0.8 compared to PCR results; however, only 39% (27/70) strain was positive based on values of rdotp > 0.95 and library dotp > 0.8. For LIAQLGGPGGVTAFAR (CTX-M-9), 100% (70/70) strains were positive for library dotp > 0.8 compared to PCR results; however, only 81% (57/70) strains were positive based on values of rdotp > 0.95 and library dotp > 0.8. This phenomenon is partly due to the existence of variants.

Overall, directly identifying peptides *via* LC-MS/MS provides a new approach to detect β-lactamases. Our results showed the great potential of the rapid extraction and detection method in the detection of β-lactamases. Started by picking the colony, peptides were obtained by rapid ultrasonic lysis and rapid digestion (even in a microwave oven for 5 min), and then results were acquired by rapid LC-MS/MS analysis (30 min) and data processing. The application of targeted proteomics using single bacterial colonies and even blood samples in the future, will enable early clinical diagnostics and early treatment. However, there are several problems that need to be addressed. For example, at present, drug-resistant enzymes are named and classified based on gene sequence; however, the protein sequence is more decisive in determining function. Additionally, there is a lack of protein subtypes and classification based on amino acid sequences. Therefore, it is difficult to identify specific peptides for one group or one subtype. Variants are ubiquitous in β-lactamase enzymes such as $bla_{\mathrm{CTX-M}}$. The separation of variants with a similar m/z and RT through regular LC-MS/MS remains a large challenge.

In conclusion, the rapid and accurate identification of β-lactamases is of great significance to clinical diagnostics and treatment. This study describes a rapid extraction and detection workflow for widespread β-lactamases, including KPC, IMP,

VIM, NDM, OXA, CMY, CTX-M, and TEM using single colonies of *Enterobacteriaceae* strains. PRM targeted proteomics was proven to be a promising approach for the detection of drug-resistant enzymes.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

XFY and YL: conceptualization and funding acquisition. YL, XH, XW, and JP: methodology. YL and GL: software. YL: validation, formal analysis, writing—original draft preparation, and supervision. XH: resources. XYY, XFY, and CL: writing—review and editing. GL: visualization. XFY: project administration. All authors have read and agreed to the published version of the manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (32141003, 81803593, and 81621064), the CAMS Initiative for Innovative Medicine (2016-I2M-3-014), the National Mega-project for Innovative Drugs (2019ZX09721001), the MOST/MOF Funds for the National Science and Technology Resource Sharing Service Platform, China (2019-194-NPRC), and the Fundamental Research Funds for the Central Universities (3332018094).

ACKNOWLEDGMENTS

We are very grateful to Dr. Hanxing Cheng and Dr. Yingzi Qi from Thermo Fisher Scientific for the help in PRM method development.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.784628/full#supplementary-material

REFERENCES

Boutal, H., Naas, T., Devilliers, K., Oueslati, S., Dortet, L., Bernabeu, S., et al. (2017). Development and validation of a lateral flow immunoassay for rapid detection of NDM-Producing *Enterobacteriaceae*. *J. Clin. Microbiol.* 55, 2018–2029. doi: 10.1128/JCM.00248-17

Bush, K., and Bradford, P. A. (2016). β-Lactams and β-Lactamase inhibitors: an overview. *Cold Spring Harb. Perspect. Med.* 6:a025247. doi: 10.1101/cshperspect.a025247

Bush, K., and Jacoby, G. A. (2010). Updated functional classification of beta-lactamases. *Antimicrob. Agents Chemother.* 54, 969–976. doi: 10.1128/AAC.01009-09

Canton, R., Gonzalez-Alba, J. M., and Galan, J. C. (2012). CTX-M Enzymes: Origin and Diffusion. *Front. Microbiol.* 3:110. doi: 10.3389/fmicb.2012.00110

Chang, C. J., Lin, J. H., Chang, K. C., Lai, M. J., Rohini, R., and Hu, A. (2013). Diagnosis of β-Lactam resistance in *Acinetobacter baumannii* using shotgun proteomics and LC-Nano-Electrospray ionization ion trap mass spectrometry. *Anal. Chem.* 85, 2802–2808. doi: 10.1021/ac303326a

Dallenne, C., Costa, A. D., Decré, D., and Favier, C. (2010). Development of a set of multiplex PCR assays for the detection of genes encoding important beta-lactamases in *Enterobacteriaceae*. *J. Antimicrob. Chemother.* 65, 490–495. doi: 10.1093/jac/dkp498

Duin, V. D. (2017). Carbapenem-resistant *Enterobacteriaceae*: What we know and what we need to know. *Virulence* 8, 379–382. doi: 10.1080/21505594.2017.1306621

Eliopoulos, G. M., and Karen, B. (2001). New beta-lactamases in gram-negative bacteria: diversity and impact on the selection of antimicrobial therapy. *Clin. Infect. Dis.* 32, 1085–1089. doi: 10.1086/319610

Espadas, G., Borràs, E., Chiva, C., and Sabidó, E. (2017). Evaluation of different peptide fragmentation types and mass analyzers in data-dependent methods using an Orbitrap Fusion Lumos Tribrid mass spectrometer. *Proteomics* 17:1600416. doi: 10.1002/pmic.201600416

Fleurbaaij, F., Goessens, W., van Leeuwen, H. C., Kraakman, M. E. M., Bernards, S. T., Hensbergen, P. J., et al. (2017). Direct detection of extended-spectrum beta-lactamases (CTX-M) from blood cultures by LC-MS/MS bottom-up proteomics. *Eur. J. Clin. Microbiol. Infect. Dis.* 36, 1621–1628. doi: 10.1007/s10096-017-2975-y

Fleurbaaij, F., Heemskerk, A. A., Russcher, A., Klychnikov, O. I., Deelder, A. M., Mayboroda, O. A., et al. (2014). Capillary-electrophoresis mass spectrometry for the detection of carbapenemases in (multi-)drug-resistant Gram-negative bacteria. *Anal. Chem.* 86, 9154–9161. doi: 10.1021/ac502049p

Foudraine, D. E., Dekker, L. J. M., Strepis, N., Bexkens, M. L., and Goessens, W. H. F. (2019). Accurate detection of the four most prevalent carbapenemases in *E. coli* and *K. pneumoniae* by high-resolution mass spectrometry. *Front. Microbiol.* 10:2760. doi: 10.3389/fmicb.2019.02760

Hleba, L., Hlebova, M., Kovacik, A., Cubon, J., and Medo, J. (2021). Carbapenemase producing *Klebsiella pneumoniae* (KPC): What is the best MALDI-TOF MS Detection Method. *Antibiotics* 10:1549. doi: 10.3390/antibiotics10121549

Honghui, W., Drake, S. K., Chen, Y., Marjan, G., Margaret, T., Rosenberg, A. Z., et al. (2016). A novel peptidomic approach to strain typing of clinical *Acinetobacter baumannii* isolates using mass spectrometry. *Clin. Chem.* 62, 866–875.

Iovleva, A., and Doi, Y. (2017). Carbapenem-Resistant *Enterobacteriaceae*. *Clin. Lab. Med.* 37, 303–315. doi: 10.1016/j.cll.2017.01.005

Kwon, T., Jung, Y. H., Lee, S., Yun, M.-r., Kim, W., Kim, D.-W. (2016). Comparative genomic analysis of *Klebsiella pneumoniae* subsp. *pneumoniae* KP617 and PittNDM01, NUHL24835, and ATCC BAA-2146 reveals unique evolutionary history of this strain. *Gut Pathog.* 8, 1–16. doi: 10.1186/s13099-016-0117-1

MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., et al. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26, 966–968. doi: 10.1093/bioinformatics/btq054

Navin, R. (2015). Parallel reaction monitoring: a targeted experiment performed using high resolution and high mass accuracy mass spectrometry. *Int. J. Mol. Sci.* 16, 28566–28581. doi: 10.3390/ijms161226120

Philippon, A., Arlet, G., and Jacoby, G. A. (2002). Plasmid-determined AmpC-type beta-lactamases. *Antimicrob. Agents Chemother.* 46, 1–11. doi: 10.1128/aac.46.1.1-11.2002

Pitout, J. D., Nordmann, P., Laupland, K. B., and Poirel, L. (2005). Emergence of *Enterobacteriaceae* producing extended-spectrum beta-lactamases (ESBLs) in the community. *J. Antimicrob. Chemother.* 56, 52–59. doi: 10.1093/jac/dki166

Rodríguez-Baño, J., Gutiérrez-Gutiérrez, B., Machuca, I., and Pascual, A. (2018). Treatment of Infections Caused by Extended-Spectrum-Beta- Lactamase-, AmpC-, and Carbapenemase-Producing *Enterobacteriaceae*. *Clin. Microbiol. Rev.* 31:e00079-17. doi: 10.1128/CMR.00079-17

Stopfer, L. E., Flower, C. T., Gajadhar, A. S., Patel, B., Gallien, S., Lopez-Ferrer, D., et al. (2021). High-density, targeted monitoring of tyrosine phosphorylation reveals activated signaling networks in human tumors. *Cancer Res*. 81, 2495–2509. doi: 10.1158/0008-5472.CAN-20-3804

Strich, J. R., Wang, H., Cisse, O. H., Youn, J. H., Drake, S. K., Chen, Y., et al. (2019). Identification of the OXA-48 carbapenemase family by use of tryptic peptides and liquid Chromatography-Tandem mass spectrometry. *J. Clin. Microbiol.* 57:e01240-18. doi: 10.1128/JCM.01240-18

Suh, J. H., Makarova, A. M., Gomez, J. M., Paul, L. A., and Saba, J. D. (2017). An LC/MS/MS method for quantitation of chemopreventive sphingadienes in food products and biological samples. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 106, 292–299. doi: 10.1016/j.jchromb.2017.07.040

Tsai, Y. M., Wang, S., Chiu, H. C., Kao, C. Y., and Wen, L. L. (2020). Combination of modified carbapenem inactivation method (mCIM) and EDTA-CIM (eCIM) for phenotypic detection of carbapenemase-producing *Enterobacteriaceae*. *BMC Microbiol.* 20:315. doi: 10.1186/s12866-020-02010-3

Vasoo, S., Cunningham, S. A., Kohner, P. C., Simner, P. J., Mandrekar, J. N., Lolans, K., et al. (2013). Comparison of a novel, rapid chromogenic biochemical assay, the carba NP Test, with the modified hodge test for detection of Carbapenemase-Producing Gram-Negative Bacilli. *J. Clin. Microbiol.* 51, 3097–3101. doi: 10.1128/JCM.00965-13

Wang, H., Drake, S. K., Youn, J. H., Rosenberg, A. Z., Chen, Y., Gucek, M., et al. (2017). Peptide markers for rapid detection of KPC Carbapenemase by LC-MS/MS. *Sci. Rep.* 7:2531. doi: 10.1038/s41598-017-02749-2

Wang, H., Strich, J. R., Drake, S. K., Chen, Y., and Suffredini, A. F. (2019). Rapid Identification of New Delhi Metallo-β-Lactamase (NDM) Using Tryptic Peptides and LC-MS/MS. *Antimicrob. Agents Chemother.* 63:e00461-19. doi: 10.1128/AAC.00461-19

WHO (2017). *WHO Publishes List of Bacteria for Which New Antibiotics Are Urgently Needed*. Geneva: WHO.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 15 September 2022
DOI 10.3389/fmicb.2022.982650

OPEN ACCESS

EDITED BY
Karsten Becker,
University Medicine Greifswald,
Germany

REVIEWED BY
Marina Oviaño,
A Coruña University Hospital Complex (CHUAC), Spain
Jeong Hwan Shin,
Inje University Busan Paik Hospital,
South Korea

*CORRESPONDENCE
Alexia Verroken
Alexia.verroken@uclouvain.be

SPECIALTY SECTION
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

RECEIVED 30 June 2022
ACCEPTED 26 August 2022
PUBLISHED 15 September 2022

CITATION
Verroken A, Hajji C, Bressant F, Couvreur J, Anantharajah A and Rodriguez-Villalobos H (2022) Performance evaluation of the FAST™ System and the FAST-PBC Prep™ cartridges for speeded-up positive blood culture testing.
Front. Microbiol. 13:982650.
doi: 10.3389/fmicb.2022.982650

Performance evaluation of the FAST™ System and the FAST-PBC Prep™ cartridges for speeded-up positive blood culture testing

Alexia Verroken*, Chaima Hajji, Florian Bressant, Jonathan Couvreur, Ahalieyah Anantharajah and Hector Rodriguez-Villalobos

Department of Microbiology, Cliniques Universitaires Saint-Luc, Université Catholique de Louvain, Brussels, Belgium

Objectives: As time to appropriate antimicrobial therapy is major to reduce sepsis mortality, there is great interest in the development of tools for direct identification (ID) and antimicrobial susceptibility testing (AST) of positive blood cultures (PBC). Very recently, the FAST™ System (Qvella) has been developed to isolate and concentrate microorganisms directly from PBCs, resulting in the recovery of a Liquid Colony™ (LC) within 30 min. The LC can be used as equivalent of an overnight subcultured colony for downstream testing. We aimed to evaluate the performances of the FAST™ System and FAST-PBC Prep™ cartridges by testing the resulting LC for direct ID, AST and rapid resistance detection.

Materials and methods: Prospectively, FAST™ System testing was carried out on each patient's first PBC with a monomicrobial Gram-stain result. In the second arm of the study, FAST™ System testing was carried out on blood cultures spiked with multidrug-resistant bacteria. Downstream testing using the LC included MALDI-TOF MS ID with the Bruker Biotyper® smart system, rapid resistance detection testing including the Abbott Diagnostics Clearview™ PBP2a SA Culture Colony Test (PBP2a) and the Bio-Rad βLACTA™ Test (βLT). AST was performed using the Becton Dickinson Phoenix™ System or by Bio-Rad disk diffusion using filter paper disk following EUCAST 2020 breakpoint criteria.

Results: FAST™ System testing was completed on 198 prospective PBCs and 80 spiked blood cultures. After exclusion of polymicrobial blood cultures, performance evaluation compared with standard of care results was carried out on 266 PBCs. Concordant, erroneous and no ID results included 238/266 (89.5%), 1/266 (0.4%), 27/266 (10.2%) PBCs, respectively. Sensitivity and specificity for PBP2a were 100% (10/10) and 75% (15/20), respectively. Sensitivity and specificity for βLT were 95.8% (23/24) and 100% (42/42), respectively. Categorical agreement for all 160 tested strains

was 98% (2299/2346) with 1.2% (8/657) very major errors and 0.7% (10/1347) major errors.

Conclusion: FAST™ System testing is a reliable approach for direct downstream testing of PBCs including MALDI-TOF MS ID, BD Phoenix™ and Bio-Rad disk diffusion AST as well as rapid resistance testing assays. Next steps include optimal integration of the FAST™ System in the PBC workflow with a view toward clinical studies.

KEYWORDS

Qvella, FAST™ System, positive blood cultures, direct MALDI-TOF MS, bacteremia, direct antimicrobial susceptibility testing, rapid resistance detection testing

Introduction

Sepsis remains a worldwide cause of morbidity and mortality with a reported 49 million cases and an approximately 11 million avoidable deaths per year (World Health Organization, 2020). As time to appropriate antimicrobial therapy is a major factor to reduce sepsis mortality, there is a great interest in the development of tools for rapid identification (ID) and antimicrobial susceptibility testing (AST) of positive blood cultures (PBCs).

Direct ID from a PBC bottle is commonly applied in clinical microbiology laboratories either by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) or by molecular techniques. These approaches result in satisfactory analytical performances with a time saving of more than 24 h compared with overnight subculture ID results (Payne et al., 2018; Ruiz-Aragón et al., 2018). However, in this current era of increasing multi-drug resistance, ID results are frequently insufficient to decide on an optimal antimicrobial treatment and the use of rapid AST approaches remains more than ever essential. Manual methods, including cleaning, washing and concentrating microorganisms directly from the PBC to obtain a pellet for direct AST, have been used historically and studies have shown an overall categorical agreement above 90% providing results one day earlier compared to AST from subculture (Maelegheer and Nulens, 2017; Hogan et al., 2019; Infante et al., 2021). Furthermore, several commercial rapid AST systems relying on cellular imaging or turbidity measurements at consecutive points of time, have been developed and provide AST results for defined antibiotic within 5–7 h. Overall agreement with standard of care (SOC) AST has been reported between 88 and 98.7% for Gram-positive cocci and between 89.5 and 94.2% for Gram-negative bacilli (Charnot-Katsikas et al., 2017; Boland et al., 2019; Grohs et al., 2021).

Very recently, a novel approach has been developed called the FAST™ System (Qvella, Richmond Hill, Canada) designed to isolate and concentrate microorganisms directly from a PBC bottle, resulting in the recovery of a liquid colony (LC) within 30 min. Ultimately, the LC can be used as an equivalent of a solid subcultured colony, enabling use of downstream ID and AST systems available in the local clinical microbiology laboratory today. The objective of this study was to evaluate the performances of the RUO FAST™ System and the FAST-PBC Prep™ cartridges-generated LC for MALDI-TOF MS ID, manual and automated AST and rapid resistance detection testing. The study also evaluated the advantages and drawbacks of this approach in comparison with other conventional and rapid techniques currently available for the routine laboratory management of PBCs.

Materials and methods

The study was conducted at the microbiology laboratory of the Cliniques universitaires Saint-Luc – UCL, a 960-bed tertiary hospital in Brussels, Belgium. Blood specimens from patients with a suspected bloodstream infection were inoculated into blood culture bottles (BD Bactec™ Plus Aerobic, Peds Plus and Lytic Anaerobic media, Becton Dickinson, Franklin Lanes, NJ, USA) and incubated 24 h a day, 7 days a week in BD Bactec™ FX devices (BD Diagnostic Systems, Sparks, MD, USA) for a standard 5-day period. SOC management of PBCs was performed during laboratory working hours (7 AM–0 AM, 7 days per week) and detailed as the "modified workflow" in a previous publication (Verroken et al., 2016). Downstream testing was performed either on an early subculture for blood cultures detected positive between 0 and 10 AM, either directly on the PBC for blood cultures detected positive between 10 AM and 3 PM, either on an overnight subculture for blood cultures detected positive between 3 PM and 0 AM. The MALDI Biotyper® smart system (Bruker Daltonik GmbH, Bremen, Germany) was used for MALDI-TOF MS ID. Rapid resistance detection tests included the Clearview™ PBP2a SA

Culture Colony Test (PBP2a; Abbott Diagnostics, Scarborough, ME, USA) performed on all *Staphylococcus aureus* and the βLACTA™ Test (βLT; Bio-Rad, Marnes-la-Coquette, France) performed on all Enterobacterales (EB) excluding chromosomal AmpC producers. AST for staphylococci, enterococci and EB was performed using the BD Phoenix System™ (Becton Dickinson, Franklin Lakes, NJ, USA) and AST of other bacteria (streptococci and *Pseudomonas aeruginosa*) was performed by disk diffusion using filter paper disks (Bio-Rad, Marnes-la-Coquette, France). All AST results were interpretated according to the breakpoint tables for interpretation of MICs and zone diameters of The European Committee on Antimicrobial Susceptibility Testing (EUCAST) version 10.0 (valid from 2020.01.01) (The European Committee on Antimicrobial Susceptibility Testing, 2020). Detection of extended-spectrum-β-lactamases (ESBL) and derepressed AmpC β-lactamases in EB relied on combination disk testing (ESBL + AmpC Screen Kit, Rosco diagnostic, Taastrup, Denmark). Carbapenemase resistance was characterized with immunochromatographic testing using the Resist-5 OOKNV and IMP K-SeT (Coris, BioConcept, Gembloux, Belgium) enabling the detection of the OXA-163, OXA-48, KPC, NDM, VIM and IMP genes.

Study design

The study was conducted in two arms. In the initial prospective arm going over a 3-month period, FAST™ System testing was performed on the first positive-detected blood culture bottle of each patient with a monomicrobial Gram-stain result. In the second arm of the study, FAST™ System testing was performed on blood culture bottles spiked with multidrug-resistant bacteria selected from a patient strain bank stored at minus 20°C. Microorganisms and resistance profiles selected for the second arm are detailed in **Table 1**. Following three successive subcultures of the initially frozen strain, the spiking process consisted of inoculating the bottles (BD Bactec™ Plus Aerobic, Peds Plus and Lytic Anaerobic) with 10 ml of human blood from healthy volunteers and 10 μl of a 1,000-times dilution of a 0.5McF suspension from a fresh overnight subcultured isolate. Blood culture bottles were incubated in a BD Bactec™ FX device until they flagged positive.

FAST™ System

The FAST™ System testing flowchart is presented in **Figure 1**. Following the availability of a monomicrobial Gram-stain result, 2 ml of the PBC was sampled into a FAST-PBC Prep™ cartridge which was loaded into the FAST™ System. Upon a 30-min automated lysis/centrifugation process, a LC constituted of pure viable bacteria was obtained. According to the manufacturer's requirements, processing had to be performed within 16 h of blood culture positivity. Following the recovery of the LC, 1 μl was plated on a non-selective blood agar in order to verify purity on the next day. Then 1 μl was double-spotted with 1 μl of formic acid and 1 μl of matrix on a target for MALDI-TOF MS ID. Depending on the identified strain, rapid resistance detection testing was subsequently performed using 2 μl of the LC and following manufacturers' recommendations. Ultimately AST was performed requiring a variable LC volume to obtain a standardized 0.5McF suspension. Rapid resistance detection tests and AST approaches were identically applied as in SOC management. In the prospective arm, AST from LC was exclusively performed if also done through SOC workflow. AST performances of FAST™ System testing were exclusively evaluated on staphylococci, enterococci, streptococci, EB and *P. aeruginosa*. AST on other microorganisms were not assessed in this study as the number of positive samples was too small to produce valuable data.

Performance evaluation

All results obtained following FAST™ System testing were compared to SOC results considered as the reference. ID and AST discordances were verified through repeated testing from LC subculture and SOC overnight subculture. Discordant PBP2a and βLT results were evaluated upon the following day with AST results.

Identification

MALDI-TOF MS ID results from the LC were interpreted according to a defined cut-off score of 1.7 for acceptable ID to the species level. A score under cut-off led to a single repetition of LC spotting and MALDI-TOF MS testing within the same day.

Rapid resistance detection testing

Rapid resistance detection testing performances were evaluated by calculating sensitivity and specificity.

Antimicrobial susceptibility testing

AST comparison was performed in accordance with Cumitech 31A recommendations for the verification and validation of procedures in the clinical microbiology laboratory (Clark et al., 2009). AST results comparison between FAST™ System testing and SOC was expressed in categorical agreement (CA) percentage (total categorical matches/total antibiotics tested × 100). Discordances were classified into very major

TABLE 1 Microorganisms and resistance profiles selected for FAST™ System testing evaluation in the second arm of the study.

Gram-positive bacteria	*n*	Oxacilline R
Staphylococci		
Staphylococcus aureus	15	8
Staphylococcus epidermidis	7	4
Staphylococcus haemolyticus	2	2
Staphylococcus hominis	1	1

Gram-negative bacteria	*n*	Derepressed AmpC	ESBL	Carbapenemase	Chromosomic carbapenem R
Enterobacterales					
Citrobacter freundii	1	1	0	0	0
Enterobacter cloacae complex	5	0	2	2 NDM, 1 OXA-48	0
Escherichia coli	13	2	10	1 NDM	0
Enterobacter aerogenes	2	2	0	0	0
Klebsiella pneumoniae	9	0	7	0	1
Proteus mirabilis	2	0	1	0	0
Proteus vulgaris	1	0	0	0	0
Non fermenters					
Pseudomonas aeruginosa	22	0	0	5 VIM	4

ESBL, extended-spectrum beta-lactamase; R, resistance.

FIGURE 1
FAST™ System testing flowchart. AST, antimicrobial susceptibility testing; ID, identification; LC, liquid colony; PBC, positive blood culture; PBP2a, Clearview™ PBP2a SA Culture Colony Test; βLT, βLACTA Test.

errors (VME: false susceptibility with AST performed on LC), major errors (ME: false resistance with AST performed on LC) and minor errors (MinE: reference test result susceptible at increased exposure or in the area of technical uncertainty while AST performed on LC susceptible or resistant, or vice versa). The VME rate was calculated by dividing the number of VME by the number of resistant bacteria (reference method) × 100 and The ME rate was calculated by dividing the number of ME by the number of susceptible bacteria (reference method) × 100. MinE rate was calculated by dividing the number of MinE by the total number of strains tested × 100. Acceptable performance rates for CA should be ≥90%, whereas acceptable performance

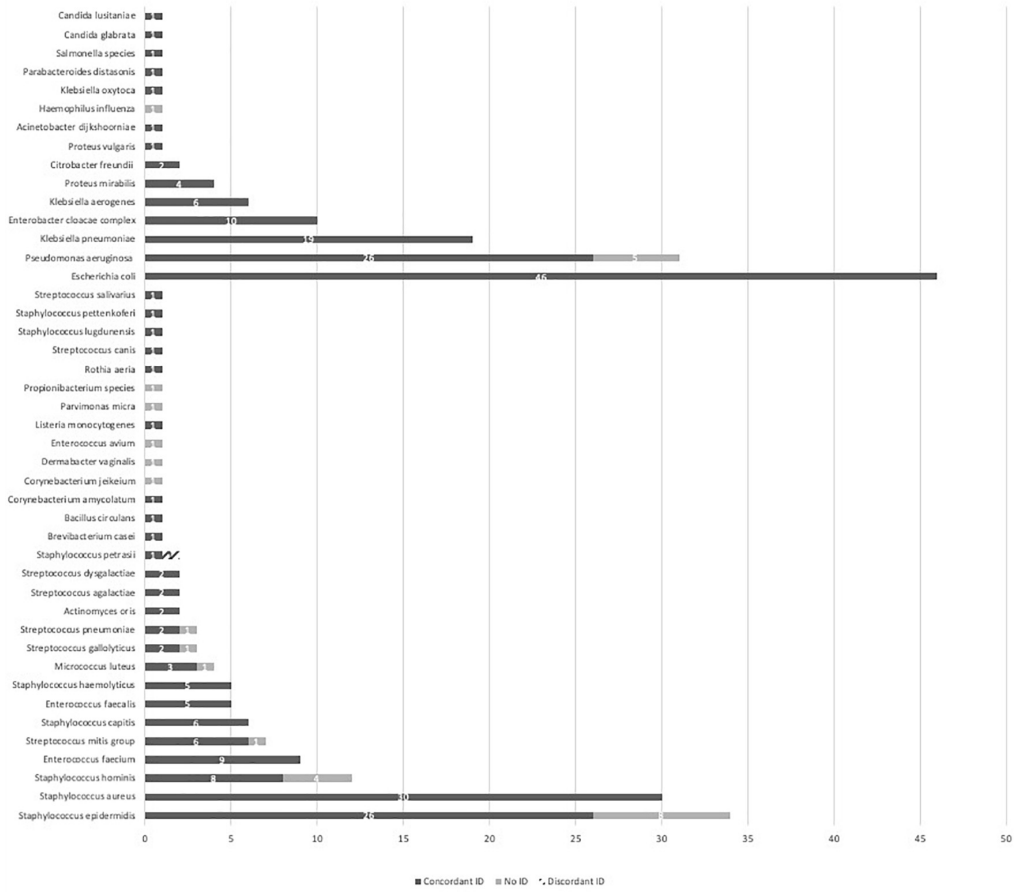


FIGURE 2
MALDI-TOF MS identification performances from a liquid colony following FAST^{TM} System testing on positive blood cultures in both study arms. ID, identification; MALDI-TOF MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry.

for the VME rate should be ≤3%. The ME rate should be ≤3%. For ME and MinE combined, the error rate should be combined ≤7%.

Results

In the prospective arm, FAST^{TM} System testing was performed on 198 patient PBCs. Ten samples were excluded from analysis because they were polymicrobial on the purity control plate on day one, 1 sample was discarded due to an instrument error and 1 sample did not have a final SOC ID. In the second arm of the study FAST^{TM} System testing was performed on 80 spiked PBC.

Identification results

Complete data are presented in **Figure 2**. Overall concordant ID was observed in 238/266 (89.5%) PBC with a mean MALDI-TOF MS score from LC testing of 2.1. Gram-positive bacteria, Gram-negative bacteria and yeast reached concordant ID results of respectively, 118/140 (84.3%), 118/124 (95.2%) and 2/2 (100%). An insufficient score resulting in the absence of ID was observed in 27/266 (10.2%) PBC. Among the non-identified PBC, 21 concerned Gram-positive bacteria including 8 *Staphylococcus epidermidis*, 4 *Staphylococcus hominis*, 3 streptococci, 1 enterococcus and 5 other strains most of the time considered as blood culture bottle contaminants. Six concerned Gram-negative bacteria including 5 *P. aeruginosa* and 1 *Haemophilus influenza*. Ultimately 1/266 (0.4%) PBC ID led to a discordant result. A *Staphylococcus petrasii* was erroneously identified as *Staphylococcus capitis* with MALDI-TOF MS from the LC with an ID score of 1.93.

Rapid resistance detection test results

PBP2a was evaluated on a total of 30 *S. aureus* including 10 methicillin-resistant *S. aureus*. Sensitivity and specificity

of PBP2a testing from a LC compared to SOC PBP2a were, respectively, 100% (10/10) and 75% (15/20). βLT was performed on 66 EB including 22 ESBL strains, one NDM-producing *Escherichia coli* and one ESBL *Klebsiella pneumoniae* with a chromosomic carbapenem resistance. Sensitivity and specificity of βLT testing from a LC compared to SOC βLT were, respectively, 95.8% (23/24) and 100% (42/42). A false negative LC βLT result on the NDM-producing *E. coli* was observed.

Antimicrobial susceptibility testing results

Automated AST was performed through both SOC testing and from a LC on a total of 127 PBC counting 54 Gram-positive strains and 73 EB. In addition, disk diffusion AST was compared between SOC testing and from a LC on 8 streptococci and 25 *P. aeruginosa*. Importantly, AST testing from LC could not be performed on 13/173 (7.5%) PBC pathogens due to insufficient LC biomass to obtain a 0.5 McF concentration.

Table 2 illustrates result comparison of automated AST on staphylococci and enterococci including 26 *S. aureus*, 15 S. coagulase negative, 8 *Enterococcus faecium* and 5 *E. faecalis*. CA, VME, ME and MinE rates were, respectively, 97.7% (650/665), 1.4% (2/144), 1.9% (9/486) and 1.3% (4/300). Moxifloxacin and tobramycin were false susceptible using the LC for 2 distinct *S. aureus* isolates. The majority of ME were found with Tobramycin 5/28 (17.9%).

Table 3 shows data comparison of automated AST on 73 EB including 42 *E. coli*, 13 *K. pneumoniae*, 6 *Enterobacter cloacae* complex, 6 *Klebsiella aerogenes*, 2 *Citrobacter freundii*, 2 *Proteus mirabilis*, 1 *Klebsiella oxytoca* and 1 *Proteus vulgaris*. CA, VME, ME and MinE rates were 97.8% (1311/1340), 1.4% (6/427), 0.1% (1/771) and 2.8% (22/788) respectively. VME were observed in 6 distinct EB strains and 5 different antibiotics. The single ME was seen with gentamicin and *P. vulgaris*.

Disk diffusion AST performances using the LC colony were evaluated on 8 streptococci including 3 *Streptococcus agalactiae*, 2 *Streptococcus dysgalactiae* and 3 *Streptococcus mitis* group. CA reached 100% with a total of 41 antibiotic combinations tested as presented in Table 4.

Ultimately AST was assessed on 25 *P. aeruginosa* resulting in a CA of 99% (297/300), no VME, no ME and 3/250 (1.2%) MinE as detailed in Table 5. Altogether AST using the LC resulted in acceptable rates according to Cumitech criteria for all evaluated automated and manual AST approaches.

Discussion

For many years, numerous laboratories have developed their own manual, in-house techniques to concentrate microorganisms from PBCs aiming to perform direct downstream testing. Throughout time, automated AST systems were evaluated for their combined performances of ID and AST directly from a PBC. In 2012, Gherardi et al. performed a comparative evaluation of the Vitek™ (bioMérieux, Marcy l'Etoile, France) and BD Phoenix™ systems for rapid ID and AST, from a standardized bacterial pellet obtained through various centrifugation steps of the PBC (Gherardi et al., 2012). Altogether, 100 and 92.3% of the Gram-negative isolates and 75 and 43.75% of the Gram-positive isolates showed concordant ID between the direct and standard methods with Vitek™ and BD Phoenix™, respectively. Additionally, AST CA of 98.7 and 99% in Gram-negative and of 96.2 and 99.5% in Gram-positive isolates with Vitek™ and BD Phoenix™, respectively, were observed. Historically multiple laboratories have performed similar evaluations yet with reduced hands-on time of pellet preparation steps. Reported performances of direct MALDI-TOF MS ID for Gram-negative bacteria exceeded 95%. However, a much lower SOC concordance rate of 79% was reached for Gram-positive bacteria (Maelegheer and Nulens, 2017; Hogan et al., 2019; Infante et al., 2021). In addition, these studies reported a CA that varied according to the evaluated AST automate between 92.9 and 98.9% using the PBC-derived pellet. FAST™ System testing results from our study were at least equal or surpassed the latter with an ID

TABLE 2 Results from automated AST (BD Phoenix™ System) of 54 staphylococci and enterococci from a liquid colony following FAST™ System testing on positive blood cultures in both study arms combined.

Staphylococci and enterococci *n* = 54

Antibiotic	CA (%)	VME (%)	ME (%)	MinE (%)
Amikacin	40/41 (97.6)	0/7 (0)	1/34 (2.9)	0/41 (0)
Ampicillin	13/13 (100)	0/8 (0)	0/5 (0)	0/13 (0)
Cefoxitin	26/26 (100)	0/9 (0)	0/17 (0)	NA
Ciprofloxacin	41/41 (100)	0/12 (0)	0/0	0/41 (0)
Clindamycin	41/41 (100)	0/12 (0)	0/29 (0)	0/41 (0)
Erythromycin	41/41 (100)	0/21 (0)	0/20 (0)	0/41 (0)
Gentamicin	39/41 (95.1)	0/8 (0)	2/33 (6.1)	NA
Gentamicin (HLR)	13/13 (100)	0/4 (0)	0/9 (0)	NA
Linezolid	54/54 (100)	0/0	0/54 (0)	NA
Moxifloxacin	40/41 (97.6)	1/12 (8.3)	0/29 (0)	NA
Oxacillin	41/41 (100)	0/18 (0)	0/23 (0)	NA
Rifampicin	41/41 (100)	0/3 (0)	0/38 (0)	0/41 (0)
Teicoplanin	54/54 (100)	0/0	0/54 (0)	NA
Tetracycline	38/41 (92.7)	0/6 (0)	0/30 (0)	3/41 (7.3)
Trimethoprim-sulfamethoxazole	39/41 (95.1)	0/11 (0)	1/29 (3.4)	1/41 (2.4)
Tobramycin	35/41 (85.4)	1/13 (8.0)	5/28 (17.9)	NA
Vancomycin	54/54 (100)	0/0	0/54 (0)	NA
Total	650/665 (97.7)	2/144 (1.4)	9/486 (1.9)	4/300 (1.3)

CA, categorical agreement; VME, very major error; ME, major error; MinE, minor error; NA, not applicable.

SOC concordance reaching 84.3 and 95.2% for Gram-positive and Gram-negative bacteria, respectively. Considering Gram-positive bacteria, failure of ID principally concerned strains that are often considered as contaminants in PBCs. The *S. capitis* erroneously identified as *S. petrasii* was a contaminant and repeated testing from LC subculture ultimately identified a *S. capitis*. It could be supposed that the PBC was most probably a polymicrobial blood culture. The absence of ID among Gram-negative bacteria principally concerned *P. aeruginosa* strains all originating from the spiked blood culture bottles from the second arm of the study. The conservation and spiking process might have altered the quality of the strains and hereby reduced the MALDI-TOF MS ID performances from LC. No other bacteria showed similar lower ID performances in the second arm of the study versus the prospective arm. AST analysis involving disk diffusion and BD Phoenix^{TM} testing, showed excellent performances substantially outperforming Cumitech requirements for all groups of bacteria frequently identified in PBC (Clark et al., 2009). Completing the prospective arm of the study with FAST^{TM} System testing on spiked blood cultures aimed to broaden AST evaluation on multidrug-resistant bacteria and did not increase VME rates. The sporadic VME involved different antibiotics tested on distinct bacteria allowing us to conclude that no specific trend was present. Nine of the 10 ME in our study were with aminoglycosides, however, their limited use as a targeted treatment of bacteriemia downplays the clinical impact of this observation. Performance data on the FAST^{TM} System generated LC are very scarce as the approach has only been marketed very recently. Similar analysis using the FAST^{TM} System testing-generated LC reported MALDI-TOF MS ID concordances of 94% and CA with Vitek^{TM} 2 for

TABLE 3 Results from automated AST (BD Phoenix^{TM} System) of 73 Enterobacterales from a liquid colony following FAST^{TM} System testing on positive blood cultures in both study arms combined.

Enterobacterales *n* = 73

Antibiotic	CA (%)	VME (%)	ME (%)	MinE (%)
Amikacin	73/73 (100)	0/7 (0)	0/66 (0)	NA
Amoxicillin-clavulanic acid	72/73 (98.6)	1/43 (2.3)	0/30 (0)	NA
Ampicillin	73/73 (100)	0/59 (0)	0/14 (0)	NA
Cefepime	70/73 (95.9)	0/24 (0)	0/43	3/73 (4.1)
Ceftriaxone	69/73 (94.5)	2/35 (5.7)	0/37 (0)	2/73 (2.7)
Ceftazidime	73/73 (100)	0/28 (0)	0/40 (0)	0/73 (0)
Cefuroxime iv	58/58 (100)	0/27 (0)	0/0	0/58 (0)
Ciprofloxacin	67/73 (91.8)	0/29 (0)	0/37	6/73 (8.2)
Ertapenem	73/73 (100)	0/10 (0)	0/63 (0)	NA
Gentamicin	71/73 (97.3)	1/14 (7.1)	1/59 (1.7)	NA
Imipenem	71/73 (97.3)	0/3 (0)	0/66 (0)	2/73 (2.7)
Levofloxacin	71/73 (97.3)	0/28 (0)	0/39 (0)	2/73 (2.7)
Meropenem	73/73 (100)	0/3 (0)	0/67 (0)	0/73 (0)
Piperacillin	72/73 (98.6)	0/50 (0)	0/22 (0)	1/73 (1.4)
Piperacillin-tazobactam	67/73 (91.8)	1/19 (5.3)	0/50 (0)	5/73 (6.8)
Tigecycline	41/41 (100)	0/0	0/41 (0)	NA
Trimethoprim-sulfamethoxazole	71/73 (97.3)	1/30 (3.3)	0/42 (0)	1/73 (1.4)
Tobramycin	73/73 (100)	0/18 (0)	0/55 (0)	NA
Total	1311/1340 (97.8)	6/427 (1.4)	1/771 (0.1)	22/788 (2.8)

CA, categorical agreement; VME, very major error; ME, major error; MinE, minor error; NA, not applicable.

TABLE 4 Results from disk difffusion AST (filter paper disks, Bio-Rad) of 8 streptococci from a liquid colony following FAST^{TM} System testing on positive blood cultures in both study arms combined.

Streptococci *n* = 8

Antibiotic	CA (%)	VME (%)	ME (%)	MinE (%)
Benzylpenicillin	8/8 (100)	0/2 (0)	0/6 (0)	0/8 (0)
Clindamycin	8/8 (100)	0/2 (0)	0/6 (0)	0/8 (0)
Erythromycin	5/5 (100)	0/2 (0)	0/3 (0)	0/5 (0)
Minocycline	5/5 (100)	0/3 (0)	0/2 (0)	0/5 (0)
Moxifloxacin	5/5 (100)	0/1 (0)	0/4 (0)	NA
Rifampicin	5/5 (100)	0/0	0/5 (0)	0/5 (0)
Trimethoprim-sulfamethoxazole	5/5 (100)	0/0	0/5 (0)	0/5 (0)
Total	41/41 (100)	0/10 (0)	0/31 (0)	0/36 (0)

CA, categorical agreement; VME, very major error; ME, major error; MinE, minor error; NA, not applicable.

TABLE 5 Results from disk difffusion AST (filter paper disks, Bio-Rad) of 25 *Pseudomonas aeruginosa* from a liquid colony following FAST^{TM} System testing on positive blood cultures in both study arms combined.

***P. aeruginosa* *n* = 25**

Antibiotic	CA (%)	VME (%)	ME (%)	MinE (%)
Amikacin	25/25 (100)	0/4 (0)	0/21 (0)	NA
Aztreonam	23/25 (92)	0/2 (0)	0/0	2/25 (8)
Cefepime	24/25 (96)	0/6 (0)	0/0	1/25 (4)
Ceftazidime	25/25 (100)	0/7 (0)	0/0	0/25 (0)
Ciprofloxacin	25/25 (100)	0/7 (0)	0/0	0/25 (0)
Imipenem	25/25 (100)	0/8 (0)	0/0	0/25 (0)
Meropenem	25/25 (100)	0/7 (0)	0/17 (0)	0/25 (0)
Piperacillin	25/25 (100)	0/8 (0)	0/0	0/25 (0)
Piperacillin-tazobactam	25/25 (100)	0/8 (0)	0/0	0/25 (0)
Ticarcillin	25/25 (100)	0/7 (0)	0/0	0/25 (0)
Ticarcillin-clavulanic avid	25/25 (100)	0/8 (0)	0/0	0/25 (0)
Tobramycin	25/25 (100)	0/4 (0)	0/21 (0)	NA
Total	297/300 (99)	0/76 (0)	0/59 (0)	3/250 (1.2)

CA, categorical agreement; VME, very major error; ME, major error; MinE, minor error; NA, not applicable.

Gram-positive and Gram-negative bacteria of 97.4 and 98.5%, respectively (Grinberg et al., 2022).

An added value of this study was the evaluation of rapid resistance detection testing performed on the LC. PBP2a testing using the LC showed suboptimal results as 5/20 methicillin-susceptible *S. aureus* (MSSA) PBC were erroneously tested positive. Initially observed in the prospective arm with 3 false positive PBP2a results on a total of 14 MSSA strains, the issue was confirmed in the second arm of the study with 2 false positive PBP2a results among 7 MSSA. All PBP2a testing results in the SOC workflow were in concordance with AST results. To our knowledge, similar observations were not reported by others. Satisfying sensitivity and specificity performances were reported by Kolesnik-Goldmann et al. who evaluated PBP2A testing on 4–6 h *S. aureus* subcultures (Kolesnik-Goldmann et al., 2021). Similarly, Defourny et al. reported a 100% sensitivity and specificity of PBP2A directly performed on a home-made PBC pellet (Defourny et al., 2014). A likely hypothesis that could explain the poor specificity results of this study is a cross-reaction of one of the reagents included in the FAST-PBC Prep™ cartridge with the recombinant monoclonal antibody fragments of the test membrane. While awaiting the conclusions of additional analyses currently performed at Qvella's, extreme vigilance is recommended in the interpretation of negative PBP2a results obtained from a LC. On the contrary, βLT study results showed very satisfying performances and hereby confirmed previous results on EB excluding AmpC chromosomal producers (Prod'hom et al., 2015).

Despite a monomicrobial Gram stain, 10 PBC of the prospective arm were ultimately excluded from data analysis as they led to a polymicrobial growth on the subcultured purity control plate on day 1. Nevertheless, when used in routine FAST™ system test results would have already been made available prior to the polymicrobial detection. Considering the 10 PBC, 3 PBC tested with the FAST™ system did not lead to any ID results and 7 PBC enabled ID of 1 out of the 2 strains. There were no erroneous IDs. The lack of ability of MALDI-TOF MS to detect all micro-organisms in mixed cultures through direct ID is well-known and reported in various publications (Verroken et al., 2015). Culturing a purity plate remains therefore essential.

While awaiting the availability of innovative sepsis diagnostic tools skipping blood incubation, a plethora of methods to speed up results from PBCs are existing (Dubourg et al., 2018; Peker et al., 2018). The FAST™ System using a LC belongs to the category of techniques aiming at the rapid production of a "microorganism pellet or suspension" with the same characteristics as an overnight subcultured colony allowing immediate downstream testing. This distinct approach combines several advantages. First of all, ID can be performed with MALDI-TOF MS enabling access to a nearly "universal" bank of bacterial fingerprints in contrast with direct molecular methods giving access to a restricted panel of strain IDs with previous publications reporting 85.2–89.1% coverage of PBC organisms (Verroken et al., 2019; Ullberg and Özneci, 2020). Additionally, no sacrifices have to be made on the selection of tested antibiotics for AST. Recent commercial AST tools designed for direct testing on PBC use a restricted number of antibiotics and don't allow the selection of a panel in accordance with the local resistance strain epidemiology of each medical center (Charnot-Katsikas et al., 2017; Boland et al., 2019; Grohs et al., 2021). FAST™ System testing enables the use of SOC well-known largely validated antimicrobial approaches including disk diffusion or automated AST, initiated from a standardized inoculum hereby enhancing the accuracy of the AST analytical performances. A procedural benefit of the FAST™ System is the complete automation of the approach with human intervention being limited to the sampling of 2 ml PBC into a cartridge and minimal run start time. Maximum number of samples that can be run at once are 2 PBC tests. The short 30-min turn-around-time for the creation of the LC enables a testing flow of approximately 20 PBC/day knowing that only the first PBC of an episode ultimately requires speeded-up testing. In a clinical microbiology laboratory that does not have a night shift, FAST™ System testing could be performed on blood cultures detected positive until approximately 5 PM ensuring ID on the same day and AST results the following day. Eventually the use of a commercial approach for SOC microbiology testing facilitates the process toward accreditation as only a method verification is required in contrast to usually fastidious and broad method validations for in-house approaches.

Originally the design of our study was not thought to integrate time-to-result measurements compared to SOC. FAST™ System testing was performed throughout the workday right upon PBC detection yet depending on the availability of the device. Subsequently, downstream testing was done in batches in the late morning and late afternoon. We can therefore affirm that ID and rapid resistance detection test results were available on the day the blood culture flagged positive within a maximum time period of 5 h following FAST™ System testing while the first AST results were available in the late evening or during the night. These turn-around-times are aligned and even beyond targeted SOC timeframe objectives of PBC management in clinical microbiology laboratories. How the FAST™ System can be integrated in the SOC PBC workflow to optimize its time-saving advantage will definitely vary from one laboratory to another depending on laboratories' working hours, currently-used downstream tests but also the existing interactions with clinicians as well as ongoing antimicrobial stewardship programs. Important to note is the recent overview presented by Lamy et al. illustrating that progress in PBC management should be based on a bundle approach joining rapid diagnostics with pre-analytical improvements, optimized microbiologistics and customized result communication (Lamy et al., 2020).

Eventually a new technology is only fully effective upon the demonstration of its clinical impact. In a recent review, Banerjee et al. concluded that rapid AST methods on PBC can shorten time to optimal treatment and improve antibiotic stewardship, however, they did not demonstrate significant reductions in mortality, length of stay or adverse effects (Banerjee and Humphries, 2021). It is therefore essential to set up well-designed clinical randomized controlled trials targeting specific patient populations and promoting clinicians' interactions to value the real impact of FAST™ System testing.

In conclusion, our results show that generating a LC through FAST™ System testing is a reliable approach to speed up downstream testing of a PBC with satisfying performances considering MALDI-TOF MS ID, disk diffusion and BD Phoenix™ AST. Next steps include its optimal integration into SOC PBC routine workflow and the set-up of effective clinical and economical studies.

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Comité d'Éthique Hospitalo-Facultaire, Cliniques Universitaires Saint-Luc. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

AV designed the study and analyzed all the results and wrote the manuscript. CH, FB, and JC performed the experiences. AA and HR-V provided critical feedback and contributed to the final version of the manuscript. All authors approved the submitted version.

Funding

This study was supported by Qvella (Ontario, Canada) who supplied the FAST™ System as well as the FAST™-PBC Prep cartridges to perform the study. The funding source had no influence on evaluation and result interpretation.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Banerjee, R., and Humphries, R. (2021). Rapid antimicrobial susceptibility testing methods for blood cultures and their clinical impact. *Front. Med.* 8:635831. doi: 10.3389/fmed.2021.635831

Boland, L., Streel, C., De Wolf, H., Rodriguez, H., and Verroken, A. (2019). Rapid antimicrobial susceptibility testing on positive blood cultures through an innovative light scattering technology: Performances and turnaround time evaluation. *BMC Infect. Dis.* 19:989. doi: 10.1186/s12879-019-4623-x

Charnot-Katsikas, A., Tesic, V., Love, N., Hill, B., Bethel, C., Boonlayangoor, S., et al. (2017). Use of the accelerate Pheno system for identification and antimicrobial susceptibility testing of pathogens in positive blood cultures and impact on time to results and workflow. *J. Clin. Microbiol.* 56:e01166-17. doi: 10.1128/JCM.01166-17

Clark, R. B., Loeffelholz, M. J., and Tibbetts, R. J. (2009). *Verification and validation of procedures in the clinical microbiology laboratory: Cumitech 31A: American society for microbiology*. Washington, DC: ASM Press.

Defourny, L., Simar, J., and Verroken, A. (2014). "Evaluation of BinaxNOW Staphylococcus aureus and PBP2a Culture Colony Test for rapid identification and resistance detection of Staphylococcus aureus on positive blood culture bottles," in *Proceedings of the 44th Interscience Conference on Antimicrobial Agents and Chemotherapy*, San Francisco, CA.

Dubourg, G., Lamy, B., and Ruimy, R. (2018). Rapid phenotypic methods to improve the diagnosis of bacterial bloodstream infections: Meeting the challenge to reduce the time to result. *Clin. Microbiol. Infect.* 24, 935–943. doi: 10.1016/j.cmi.2018.03.031

Gherardi, G., Angeletti, S., Panitti, M., Pompilio, A., Di Bonaventura, G., Crea, F., et al. (2012). Comparative evaluation of the Vitek-2 Compact and Phoenix systems for rapid identification and antibiotic susceptibility testing directly from blood cultures of Gram-negative and Gram-positive isolates. *Diagn. Microbiol. Infect. Dis.* 72, 20–31. doi: 10.1016/j.diagmicrobio.2011.09.015

Grinberg, S., Schubert, S., Hochauf-Stange, K., Dalpke, A. H., and Narvaez Encalada, M. (2022). Saving time in blood culture diagnostics: A prospective evaluation of the Qvella FAST™-PBC Prep application on the FAST™ system. *J. Clin. Microbiol.* 60:e02533-21. doi: 10.1128/jcm.02533-21

Grohs, P., Picard, S., Mainardi, J. L., and Podglajen, I. (2021). Assessment of version 2.5 of QMAC-dRAST for rapid antimicrobial susceptibility testing with reduced sample-to-answer turnaround time and an integrated expert system. *Infect. Dis. Now* 51, 470–476. doi: 10.1016/j.idnow.2020.10.005

Hogan, C. A., Watz, N., Budvytiene, I., and Banaei, N. (2019). Rapid antimicrobial susceptibility testing by VITEK® 2 directly from blood cultures in patients with Gram-negative rod bacteremia. *Diagn. Microbiol. Infect. Dis.* 94, 116–121. doi: 10.1016/j.diagmicrobio.2019.01.001

Infante, A., Ortiz de la Tabla, V., Martín, C., Gázquez, G., and Buñuel, F. (2021). Rapid identification and antimicrobial susceptibility testing of Gram-negative rod on positive blood cultures using MicroScan panels. *Eur. J. Clin. Microbiol. Infect. Dis.* 40, 151–157. doi: 10.1007/s10096-020-04014-3

Kolesnik-Goldmann, N., Bodendoerfer, E., Röthlin, K., Herren, S., Imkamp, F., Marchesi, M., et al. (2021). Rapid detection of PBP2a in Staphylococci from shortly incubated subcultures of positive blood cultures by an immunochromatographic assay. *Microbiol. Spectr.* 9:e00462-21. doi: 10.1128/Spectrum.00462-21

Lamy, B., Sundqvist, M., Idelevich, E. A., and Escmid Study Group for Bloodstream Infections, Endocarditis and Sepsis [ESGBIES] (2020). Bloodstream infections - Standard and progress in pathogen diagnostics. *Clin. Microbiol. Infect.* 26, 142–150. doi: 10.1016/j.cmi.2019.11.017

Maelegheer, K., and Nulens, E. (2017). Same-day identification and antibiotic susceptibility testing on positive blood cultures: A simple and inexpensive procedure. *Eur. J. Clin. Microbiol. Infect. Dis.* 36, 681–687. doi: 10.1007/s10096-016-2849-8

Payne, M., Champagne, S., Lowe, C., Leung, V., Hinch, M., and Romney, M. G. (2018). Evaluation of the FilmArray Blood Culture Identification Panel compared to direct MALDI-TOF MS identification for rapid identification of pathogens. *J. Med. Microbiol.* 67, 1253–1256. doi: 10.1099/jmm.0.000802

Peker, N., Couto, N., Sinha, B., and Rossen, J. W. (2018). Diagnosis of bloodstream infections from positive blood cultures and directly from blood samples: Recent developments in molecular approaches. *Clin. Microbiol. Infect.* 24, 944–955. doi: 10.1016/j.cmi.2018.05.007

Prod'hom, G., Durussel, C., Blanc, D., Croxatto, A., and Greub, G. (2015). Early detection of extended-spectrum β-lactamase from blood culture positive for an *Enterobacteriaceae* using βLACTA test. *New Microbes New Infect.* 8, 1–3. doi: 10.1016/j.nmni.2015.05.007

Ruiz-Aragón, J., Ballestero-Téllez, M., Gutiérrez-Gutiérrez, B., de Cueto, M., Rodríguez-Baño, J., and Pascual, Á. (2018). Direct bacterial identification from positive blood cultures using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry: A systematic review and meta-analysis. *Enferm. Infecc. Microbiol. Clin.* 36, 484–492. doi: 10.1016/j.eimc.2017.08.012

The European Committee on Antimicrobial Susceptibility Testing, (2020). *Breakpoint tables for interpretation of MICs and zone diameters. Version 10.0, 2020.* Available online at: http://www.eucast.org/clinical_breakpoints/

Ullberg, M., and Özneci, V. (2020). Identification and antimicrobial susceptibility testing og Gram-positive and Gram-negative bacteria from positive blood cultures using the Accelerate Phenoä system. *Eur. J. Clin. Microbiol. Infect. Dis. Eur. Soc. Clin. Microbiol.* 39, 139–149. doi: 10.1007/s10096-019-13703-y

Verroken, A., Defourny, L., le Polain de Waroux, O., Belkhir, L., Laterre, P. F., Delmée, M., et al. (2016). Clinical impact of MALDI-TOF MS identification and rapid susceptibility testing on adequate antimicrobial treatment in sepsis with positive blood cultures. *PLoS One* 11:e0156299. doi: 10.1371/journal.pone.0156299

Verroken, A., Defourny, L., Lechgar, L., Magnette, A., Delmée, M., and Glupcynzki, Y. (2015). Reducing time to identification of positive blood cultures with MALDI-TOF MS analysis after a 5-h subculture. *Eur. J. Clin. Microbiol. Infect. Dis.* 34, 405–413.

Verroken, A., Despas, N., Rodriguez-Villalobos, H., and Laterre, P. F. (2019). The impact of a rapid molecular identification test on positive blood cultures from critically ill with bacteremia: A pre-post intervention study. *PLoS One* 14:e0223122. doi: 10.1371/journal.pone.0223122

World Health Organization (2020). *Sepsis.* Available online at: https://www.who.int/news-room/fact-sheets/detail/sepsis (accessed May 16, 2022).

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 23 November 2022
DOI 10.3389/fmicb.2022.1059104

OPEN ACCESS

EDITED BY
Antonella Lupetti,
University of Pisa,
Italy

REVIEWED BY
Haiquan Kang,
Affiliated Hospital of Xuzhou Medical University, China
Pawan Kumar Kanaujia,
University of Delhi,
India

*CORRESPONDENCE
Elvira R. Shaidullina
elvira.shaidullina@antibiotic.ru

SPECIALTY SECTION
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

RECEIVED 30 September 2022
ACCEPTED 07 November 2022
PUBLISHED 23 November 2022

CITATION
Shaidullina ER, Romanov AV, Skleenova EY, Sheck EA, Sukhorukova MV, Kozlov RS and Edelstein MV (2022) Detection of carbapenemase-producing Enterobacterales by means of matrix-assisted laser desorption ionization time-of-flight mass spectrometry with ertapenem susceptibility-testing disks as source of carbapenem substrate. *Front. Microbiol.* 13:1059104. doi: 10.3389/fmicb.2022.1059104

Detection of carbapenemase-producing Enterobacterales by means of matrix-assisted laser desorption ionization time-of-flight mass spectrometry with ertapenem susceptibility-testing disks as source of carbapenem substrate

Elvira R. Shaidullina*, Andrey V. Romanov, Elena Y. Skleenova, Eugene A. Sheck, Marina V. Sukhorukova, Roman S. Kozlov and Mikhail V. Edelstein

Institute of Antimicrobial Chemotherapy, Smolensk State Medical University, Smolensk, Russia

MALDI-TOF mass spectrometry has become widely used in clinical microbiology and has proved highly accurate for detection of carbapenemases in Gram-negative bacteria. However, the use of carbapenem-hydrolysis assays in routine diagnostics is hampered by the need for antibiotic substances and for making their fresh solutions each time an assay is conducted. Here, we evaluated the use of commercial antibiotic susceptibility-testing disks as source of ertapenem substrate in MALDI-TOF MS-based assay for detection of carbapenemase-producing Enterobacterales (CPE). The assay was validated on 48 CPE isolates of 8 different species expressing NDM-, VIM-, KPC- and OXA-48-type carbapenemases and exhibiting various levels of resistance to carbapenems (MIC range: 0.25–>32 mg/l), as well as on 48 carbapenemase-non-producing isolates. The assay conditions were optimized as follows: 10-µl loopful of bacterial colonies was suspended in 150 µl 0.01M Na-PBS buffer, pH 7.4, a 10 µg ertapenem susceptibility-testing disk was immersed in the suspension and incubated 3h at 35°C, after which supernatant was obtained by centrifugation and applied on a target plate with alpha-cyano-4-hydroxycinnamic acid matrix. Mass spectra were analyzed between 440 and 560m/z. Carbapenemase activity was detected in all tested CPE isolates by the appearance of m/z peaks corresponding to ertapenem hydrolysis products: $[M_h+H]^+$:494.2, $[M_h+Na]^+$:516.2, $[M_h+2Na]^+$:538.2, $[M_{h/d}+H]^+$:450.2, $[M_{h/d}+Na]^+$:472.2, and simultaneous decrease or loss of peaks of intact antibiotic: $[M+H]^+$:476.2, $[M+Na]^+$:498.1, $[M+2Na]^+$:520.1. No hydrolysis peaks or loss of intact ertapenem peaks were observed for carbapenemase-negative strains. We therefore report the development of a sensitive, specific and cost-effective MALDI-TOF MS-based assay for detection of CPE, which makes use of antibiotic disks readily available in most laboratories.

KEYWORDS

MALDI-TOF MS, resistance detection, antibiotic resistance, carbapenemases, Enterobacterales

Introduction

Global spread of carbapenemase-producing *Enterobacterales* (CPE) is one of the greatest antimicrobial resistance threats to modern healthcare (Marchaim et al., 2012; Doi and Paterson, 2015; Lee et al., 2016; Potter et al., 2016; Hsu et al., 2017). Effective detection of carbapenemases is important for infection control and antibiotic treatment of CPE infections and requires fast and accurate tests. A number of such tests have been developed for use in the clinical microbiology laboratories, including molecular and immunochromatographic tests that enable rapid targeted identification of the known carbapenemases and phenotypic carbapenem hydrolysis tests that provide alternative and complementary means of detecting any carbapenem-inactivating enzymes (Nordmann et al., 2012; van der Zwaluw et al., 2015; Kieffer et al., 2019; Feng et al., 2021). Hydrolysis tests have become widely used after the introduction of the Carba NP test in 2012 (Nordmann et al., 2012). Three years later, van der Zwaluw et al. (2015) proposed an alternative simple and low-cost Carbapenem Inactivation Method (CIM) making use of common susceptibility testing disks and indicator *Escherichia coli* strain for detecting hydrolysis of meropenem, which later on received several modifications and became extremely practical for routine use (Pierce et al., 2017; Jing et al., 2018; Muntean et al., 2018).

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) has revolutionized modern microbiology and has found many applications not only in species identification of bacteria and fungi but also in detecting resistance mechanisms to antimicrobials (Claydon et al., 1996; Holland et al., 1996; Oviaño and Rodríguez-Sánchez, 2021; Torres-Sangiao et al., 2021; Yoon and Jeong, 2021). Lately, MALDI-TOF MS has been successfully applied to the detection of various β-lactamases, in particular carbapenemases (Burckhardt and Zimmermann, 2011; Hrabák et al., 2011). Unlike other phenotypic assays, which rely upon the use of various indirect indicators of β-lactamase-mediated hydrolysis of β-lactams, MALDI-TOF MS allows direct visualization of mass-peaks corresponding to intact β-lactam substrates and their degradation products released upon hydrolysis and, therefore, may be considered as a reference method for detecting β-lactamase activity. However, most MALDI-TOF MS-based assays for carbapenemase detection described up to date required preparation of fresh solutions of carbapenems either from chemical substances or therapeutic formulations, which made them less suitable for routine diagnostics (Hrabák et al., 2011; Hoyos-Mallecot et al., 2014; Knox et al., 2014; Papagiannitsis et al., 2015; Ghebremedhin et al., 2016; Sakarikou et al., 2017). Recently, M. Oho et al. reported on the development of MALDI-TOF MS assay with imipenem susceptibility disk and zinc sulfate solution for detection of CPE (Oho et al., 2021).

Herein, we describe the development of a modified MALDI-TOF MS assay that makes use of commercial antibiotic susceptibility-testing disks with ertapenem for highly sensitive and specific detection of CPE.

Materials and methods

Bacterial isolates

A representative collection of 96 non-duplicate clinical isolates from IAC collection was used. These isolates were retrospectively collected from the national sentinel surveillance program and were previously extensively characterized for susceptibility to carbapenem antibiotics using broth microdilution method, and for carbapenemase gene content using PCR and sequencing (Kuzmenkov et al., 2021). Phenotypic expression of carbapenemases was preliminary assessed using CIM test. The isolates belonged to 11 species: *Klebsiella pneumoniae* ($n = 65$), *Escherichia coli* ($n = 10$), *Enterobacter cloacae* ($n = 5$), *Proteus mirabilis* ($n = 5$), *Serratia marcescens* ($n = 3$), *Citrobacter freundii* ($n = 2$), *Klebsiella oxytoca* ($n = 2$), *Enterobacter aerogenes* ($n = 1$), *Enterobacter asburiae* ($n = 1$), *Morganella morganii* ($n = 1$), and *Proteus vulgaris* ($n = 1$). The test collection included 48 isolates producing OXA-48-like ($n = 30$), NDM ($n = 13$), VIM ($n = 2$), KPC ($n = 1$), and combination of OXA-48-like and NDM ($n = 2$) carbapenemases, which exhibited various levels of resistance to carbapenems (MIC range: 0.25–> 32 mg/l), and 48 carbapenemase-negative isolates (Supplementary Table S1).

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry assay

Several modifications of assay conditions were tested (data not shown). These included: (i) subcultivation of test isolates on Mueller-Hinton agar with or without supplementation with 10 mM zinc sulfate; (ii) preparing suspension of test cultures in pure deionized water, 0.9% non-buffered saline solution, or sodium phosphate buffered saline (Na-PBS) with or without addition of 5% propanol-2, (iii) use of susceptibility testing disks with imipenem, 10 μg, meropenem, 10 μg, doripenem, 10 μg, or

ertapenem, 10 µg (Bio-Rad Laboratories, Marnes-la-Coquette, France); (iv) use of different incubation time (from 0.5 to 4 h).

The final optimized assay conditions were as follows: Bacterial isolates were recovered from storage at −70°C in glycerol-supplemented brain-heart infusion broth and subcultured for 18 h at 35°C on Mueller–Hinton agar (BBL MH II; Becton Dickinson, Sparks, MD). A10-µl loopful of bacterial colonies was suspended in 150 µl 0.01 M Na-PBS, pH 7.4, a 10 µg ertapenem susceptibility-testing disk (Bio-Rad Laboratories) was immersed in the suspension and incubated 3 h at 35°C. The suspension was then centrifuged at 14,000 rpm for 2 min. One microliter of supernatant was applied on a steel target plate on top of the pre-dried layer of MALDI matrix (1 µl of 2.5 µg/ml alpha-cyano-4-hydroxycinnamic acid; HCCA, Bruker Daltonik, Bremen, Germany) in 50% acetonitrile, 0.1% trifluoroacetic acid dried and overlaid with the second layer of the same matrix.

Antibiotic Calibration Standard (ACS, Bruker Daltonik) was used for external instrument calibration. Mass spectra were acquired between 440–560 m/z on a Microflex LT spectrometer with flexControl software, v3.4 (Bruker Daltonik). The acquisition parameters were set as follows: ion source 1, 18.98 kV; ion source 2, 16.25 kV; lens, 0.01 kV; laser frequency, 60 Hz; digitizer trigger level, 2,500 mV; laser attenuator offset, 28%; laser attenuator range, 30%; and laser range, 15–45%. Each spot was measured using 240 laser shots in groups of 40 shots per sampling area. The MS spectra were measured automatically in at least 3 repetitions and analyzed manually by flexAnalysis software, v3.4 (Bruker Daltonik) to identify intact vs. hydrolyzed ertapenem peaks (Table 1).

Results and discussion

In the initial stage of the study, we tested different assay conditions (as briefly described in Materials and methods) and use of common susceptibility disks with ertapenem, meropenem, doripenem, and imipenem as source carbapenem substrate in MALDI-TOF MS-based hydrolysis assay. Spontaneous degradation of all three carbapenems was assessed under assay conditions at different time intervals (1 h, 2 h, 3 h, and 4 h) without bacterial culture and with carbapenemase-negative control strain of *E. coli* ATCC25922. Imipenem showed spontaneous degradation, as was revealed by appearance of detectable MS peaks of hydrolysis products after ≥3 h incubation (Supplementary materials; Figure S3A). On the other hand, meropenem, doripenem and ertapenem were notably more stable. However, the mass-peaks of intact meropenem and doripenem and their degradation products were barely detectable or not detectable.

TABLE 1 Expected molecular masses of ionic forms of intact ertapenem and its hydrolysis products.

Intact ertapenem		Ertapenem hydrolysis products	
MW, g/mol	475.2	$[M_h+H]^+$	494.2
$[M+H]^+$	476.2	$[M_h+Na]^+$	516.2
$[M+Na]^+$	498.1	$[M_h+2Na]^+$	538.2
$[M+2Na]^+$	520.1	$[M_{h/d}+H]^+$	450.2
$[M+3Na]^+$	542.1	$[M_{h/d}+Na]^+$	472.2

M, intact molecule; M_h, hydrolyzed molecule; $M_{h/d}$, hydrolyzed and decarboxylated molecule.

In contrast, ertapenem yielded readily visible MS peaks of intact molecule ions with high signal-to-noise ratios at a concentration of 66.6 µg/ml (0.14 mM) generated from 10-µg disk in 150 µl volume of test solution. Besides, the MS peaks of ertapenem degradation products obtained after hydrolysis were also easily detectable and distinguishable from each other and from non-specific peaks (Figures 1, 2), making ertapenem the most suitable substrate for detection of carbapenemases by means of MALDI-TOF MS. Notably, other reports (Hoyos-Mallecot et al., 2014; Sakarikou et al., 2017) have also demonstrated suitability of ertapenem as substrate for MALDI-TOF MS-based carbapenemase detection.

Using the final optimized assay conditions and incubation of bacterial cultures with ertapenem disk for up to 3 h, we detected ertapenem hydrolysis by all 48 isolates producing carbapenemases of OXA-48, KPC, NDM, and VIM groups, and exhibiting variable resistance levels to carbapenems, even with MICs below the clinical resistance breakpoints indicating weak carbapenemase expression. This was evidenced by complete or partial disappearance of MS peaks at m/z 476.2, 498.1, 520.1 and 542.1, corresponding, respectively, to intact ertapenem $[M+H]^+$ and its mono- $[M+Na]^+$, di- $[M+2Na]^+$, and trisodium $[M+3Na]^+$ adducts, and by simultaneous appearance of the peak at m/z 472.2 m/z corresponding to major degradation product, a monosodium adduct of hydrolyzed and decarboxylated ertapenem molecule $[M_{h/d}+Na]^+$. In addition, the peaks at 494.2 m/z, 516.2 m/z and 538.2 m/z, corresponding, respectively, to the intermediate product (hydrolyzed ertapenem $[M_h+H]^+$ and its, mono- $[M_h+Na]^+$ and disodium $[M_h+2Na]^+$ ion forms) were observed for most CPE isolates (Figure 1; Supplementary Table S1).

Though a very small peak at 450.2 m/z likely corresponding to spontaneously hydrolyzed and decarboxylated ertapenem $[M_{h/d}+H]^+$ was present in mass spectra of carbapenemase-negative isolates, its relative intensity was significantly smaller as compared to CPE isolates. No major hydrolysis peaks or loss of intact ertapenem peaks were observed for any of the carbapenemase-negative isolates (Figure 2). The different MS peak patterns derived from CPE and non-CPE isolates allowed their unambiguous discrimination. Therefore, the sensitivity and specificity of the developed assay reached 100% for the studied panel of isolates.

Unlike the previous study (Oho et al., 2021), we did not find supplementation of culture media with zinc sulfate to improve the detection of metallo-β-lactamases (MBLs) of NDM- and VIM-type (data not show). This may be explained by the use in our study of ertapenem instead of imipenem disks or by the cultivation of isolates on Mueller-Hinton agar which contains zinc

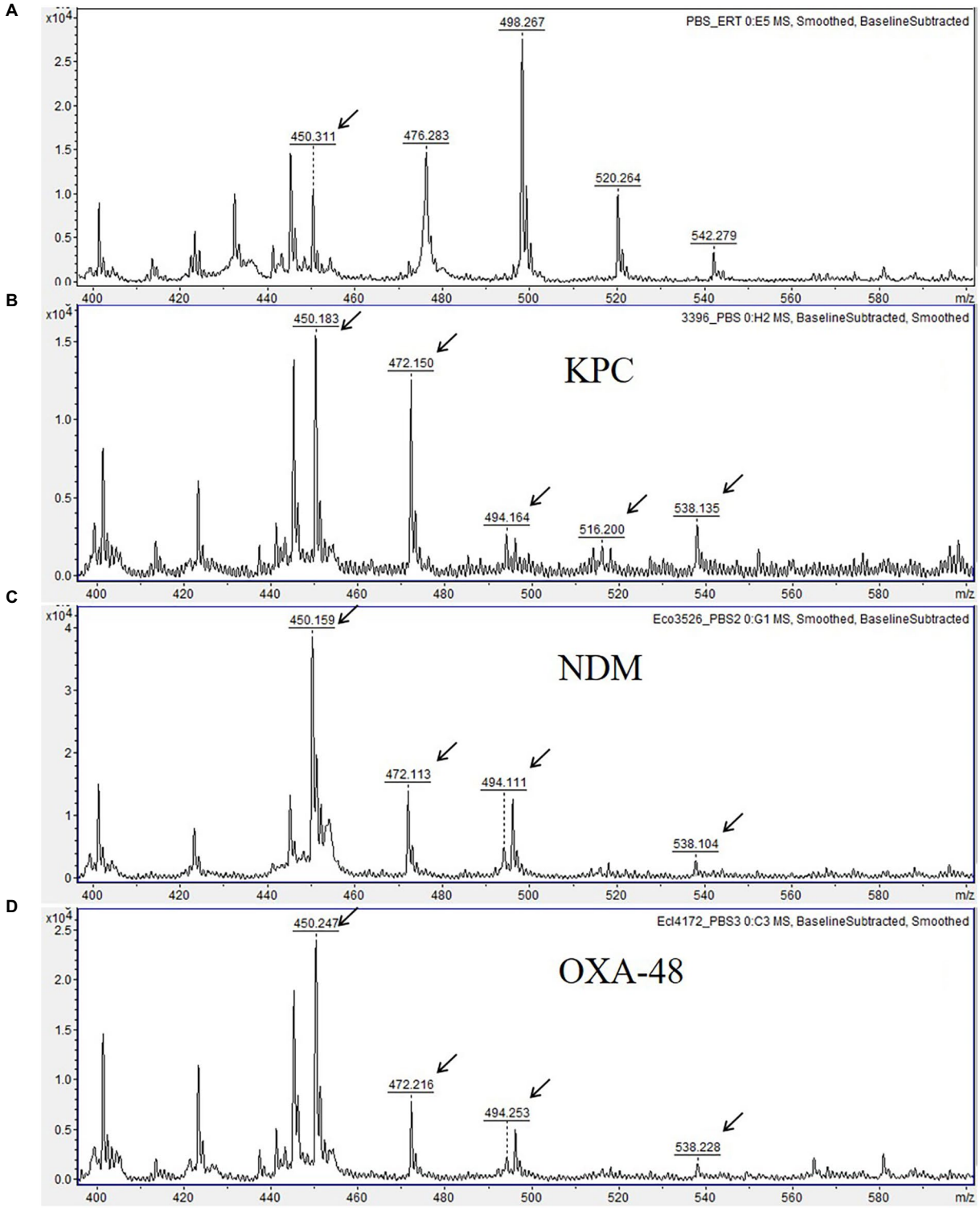


FIGURE 1
Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectra showing ertapenem hydrolysis by carbapenemase-producing Enterobacterales. **(A)** Mass spectrum of ertapenem disk after 3 h-incubation in 0.01 M PBS, pH 7.4 (negative control). **(B–D)** Mass spectra showing ertapenem hydrolysis by KPC-3, NDM-1, and OXA-48 carbapenemases. Peaks corresponding to hydrolyzed forms of ertapenem are indicated with arrows. Units on the *Y* axes represent relative intensity.

ions in concentrations sufficient for MBLs to exert their hydrolytic activity (Asempa et al., 2021). Our study, however, did not include IMP-type MBLs that were lacking in our collection, which may be considered as a limitation of the study. Further experiments may be needed to evaluate the applicability of our assay for the detection of IMP-type carbapenemases.

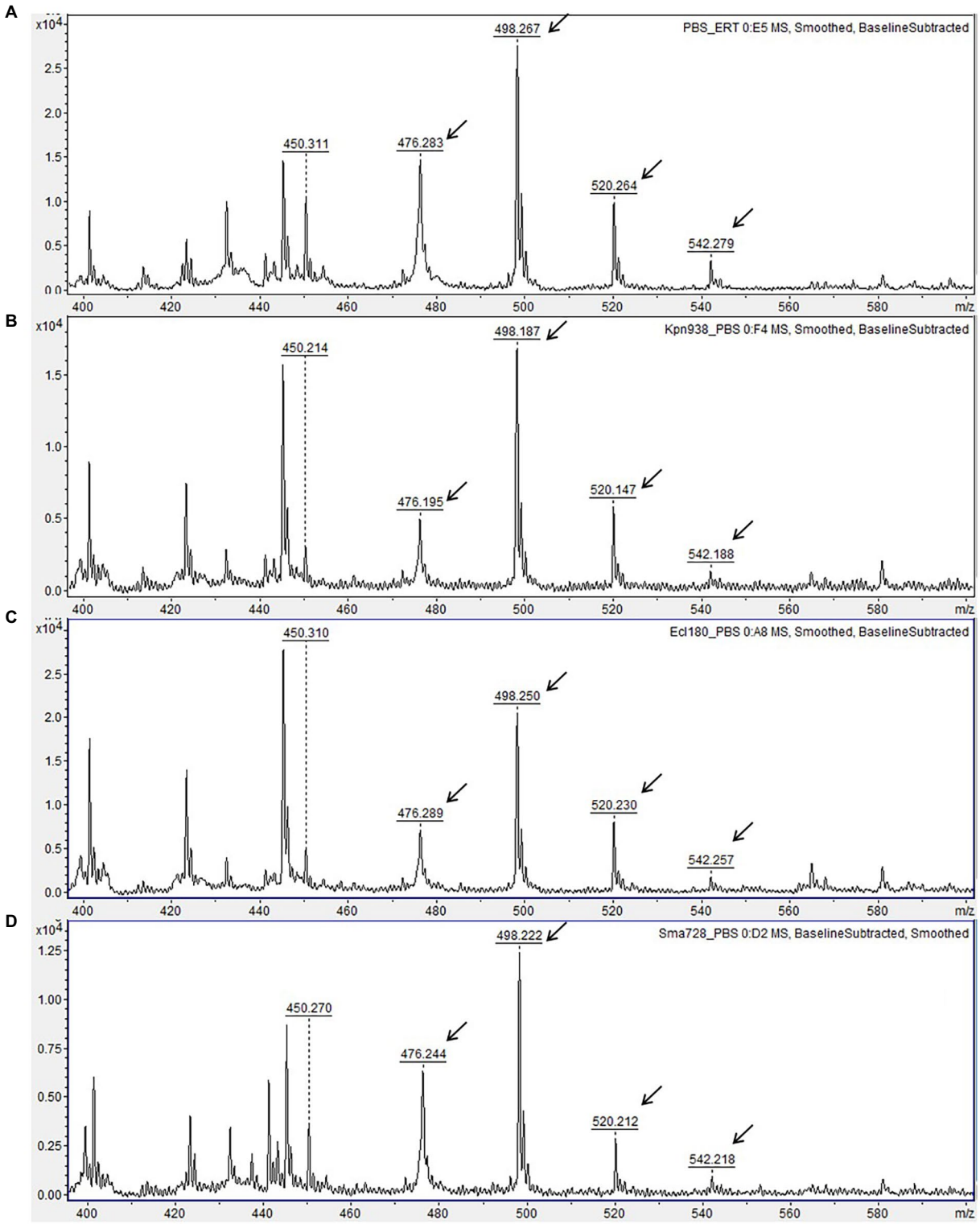


FIGURE 2
Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectra of carbapenemase-non-producing isolates. **(A)** Mass spectrum of ertapenem disk after 3 h-incubation in 0.01 M PBS, pH 7.4 (negative control). **(B–D)** Mass spectra of ertapenem after incubation with carbapenemase-non-producing isolates of different species. Peaks corresponding to non-hydrolyzed form of ertapenem are indicated with arrows. Units on the *Y* axes represent relative intensity.

Conclusion

This study describes the development of sensitive and specific MALDI-TOF MS-based assay for detection of CPE, which makes use of materials and reagents readily available in most laboratories, such as ertapenem disks used for disk-diffusion antibiotic susceptibility testing and HCCA used as MALDI matrix for species identification of bacteria and fungi. The assay does not require highly skilled personnel, and may be used in any laboratory equipped with a MALDI-TOF mass spectrometer.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

ES, ME, and RK: conceptualization. ES, AR, ElS, EuS, and MS: methodology and experimental work. ES: data validation and analysis, writing—original draft preparation. ME: writing—review and editing. RK: administration. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by Russian Federal research grant no. 1022040800475-5-1.6.2.

Acknowledgments

We would like to thank Viktor S. Kazakov for technical consultation on the assay development.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.1059104/full#supplementary-material

SUPPLEMENTARY TABLE S1
Characteristics of the isolates used in the study and presence of the marker peaks in the corresponding mass spectra.

SUPPLEMENTARY TABLE S2
Expected molecular masses of ionic forms of intact carbapenems and their hydrolysis products.

References

Asempa, T. E., Bajor, H., Mullins, J. H., Hartnett, J., and Nicolau, D. P. (2021). Evaluation of Metallo-β-lactamase susceptibility testing in a physiologic medium. *Microbiol. Spectr.* 9, e01670–e01621. doi: 10.1128/Spectrum.01670-21

Burckhardt, I., and Zimmermann, S. (2011). Using matrix-assisted laser desorption ionization-time of flight mass spectrometry to detect Carbapenem resistance within 1 to 2.5 hours. *J. Clin. Microbiol.* 49, 3321–3324. doi: 10.1128/jcm.00287-11

Claydon, M. A., Davey, S. N., Edwards-Jones, V., and Gordon, D. B. (1996). The rapid identification of intact microorganisms using mass spectrometry. *Nat. Biotechnol.* 14, 1584–1586. doi: 10.1038/nbt1196-1584

Doi, Y., and Paterson, D. L. (2015). Carbapenemase-Producing *Enterobacteriaceae*. *Semin. Respir. Crit. Care Med.* 36, 74–84. doi: 10.1055/S-0035-1544208

Feng, W., Niu, S., Chang, Y., Jia, X., Huang, S., and Yang, P. (2021). Design of Rapid Detection System for five major Carbapenemase families (bla_{KPC}, bla_{NDM}, bla_{VIM}, bla_{IMP} and $bla_{OXA\text{-}48\text{-}Like}$) by colorimetric loop-mediated isothermal amplification. *Infect. Drug Resist.* 14, 1865–1874. doi: 10.2147/IDR.S301757

Ghebremedhin, B., Halstenbach, A., Smiljanic, M., Kaase, M., and Ahmad-Nejad, P. (2016). MALDI-TOF MS based carbapenemase detection from culture isolates and from positive blood culture vials. *Ann. Clin. Microbiol. Antimicrob.* 15:5. doi: 10.1186/s12941-016-0120-x

Holland, R. D., Wilkes, J. G., Rafii, F., Sutherland, J. B., Persons, C. C., Voorhees, K. J., et al. (1996). Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 10, 1227–1232.

Hoyos-Mallecot, Y., Cabrera-Alvargonzalez, J. J., Miranda-Casas, C., Rojo-Martín, M. D., Liebana-Martos, C., and Navarro-Marí, J. M. (2014). MALDI-TOF MS, a useful instrument for differentiating metallo-β-lactamases in *Enterobacteriaceae* and *pseudomonas* spp. *Lett. Appl. Microbiol.* 58, 325–329. doi: 10.1111/lam.12203

Hrabák, J., Walková, R., Študentová, V., Chudáčková, E., and Bergerová, T. (2011). Carbapenemase activity detection by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J. Clin. Microbiol.* 49, 3222–3227. doi: 10.1128/JCM.00984-11

Hsu, L. Y., Apisarnthanarak, A., Khan, E., Suwantarat, N., Ghafur, A., and Tambyah, P. (2017). Carbapenem-resistant *Acinetobacter baumannii* and *Enterobacteriaceae* in south and Southeast Asia. *Clin. Microbiol. Rev.* 30, 1–22. doi: 10.1128/CMR.00042-16

Jing, X., Zhou, H., Min, X., Zhang, X., Yang, Q., Du, S., et al. (2018). The simplified Carbapenem inactivation method (sCIM) for simple and accurate detection of Carbapenemase-producing gram-negative bacilli. *Front. Microbiol.* 9:2391. doi: 10.3389/FMICB.2018.02391

Kieffer, N., Poirel, L., and Nordmann, P. (2019). Rapid immunochromatography-based detection of carbapenemase producers. *Infection* 47, 673–675. doi: 10.1007/S15010-019-01326-1

Knox, J., Jadhav, S., Sevior, D., Agyekum, A., Whipp, M., Waring, L., et al. (2014). Phenotypic detection of carbapenemase-producing *Enterobacteriaceae* by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry and the Carba NP test. *J. Clin. Microbiol.* 52, 4075–4077. doi: 10.1128/JCM.02121-14

Kuzmenkov, A. Y., Trushin, I. V., Vinogradova, A. G., Avramenko, A. A., Sukhorukova, M. V., Malhotra-Kumar, S., et al. (2021). AMRmap: an interactive web platform for analysis of antimicrobial resistance surveillance data in Russia. *Front. Microbiol.* 12:377. doi: 10.3389/FMICB.2021.620002

Lee, C. R., Lee, J. H., Park, K. S., Kim, Y. B., Jeong, B. C., and Lee, S. H. (2016). Global dissemination of carbapenemase-producing *Klebsiella pneumoniae*: epidemiology, genetic context, treatment options, and detection methods. *Front. Microbiol.* 7:895. doi: 10.3389/FMICB.2016.00895

Marchaim, D., Chopra, T., Bhargava, A., Bogan, C., Dhar, S., Hayakawa, K., et al. (2012). Recent exposure to antimicrobials and Carbapenem-resistant *Enterobacteriaceae*: the role of antimicrobial stewardship. *Infect. Control Hosp. Epidemiol.* 33, 817–830. doi: 10.1086/666642

Muntean, M.-M., Muntean, A.-A., Gauthier, L., Creton, E., Cotellon, G., Popa, M. I., et al. (2018). Evaluation of the rapid carbapenem inactivation method (rCIM): a phenotypic screening test for carbapenemase-producing *Enterobacteriaceae*. *J. Antimicrob. Chemother.* 73, 900–908. doi: 10.1093/jac/dkx519

Nordmann, P., Poirel, L., and Dortet, L. (2012). Rapid detection of Carbapenemase-producing *Enterobacteriaceae*. *Emerg. Infect. Dis.* 18, 1503–1507. doi: 10.3201/eid1809.120355

Oho, M., Funashima, Y., Nagasawa, Z., Miyamoto, H., and Sueoka, E. (2021). Rapid detection method of carbapenemase-producing Enterobacteriaceae by MALDI-TOF MS with imipenem/cilastatin (KB) disc and zinc sulfate solution. *J. Infect. Chemother.* 27, 205–210. doi: 10.1016/J.JIAC.2020.09.013

Oviaño, M., and Rodríguez-Sánchez, B. (2021). MALDI-TOF mass spectrometry in the 21st century clinical microbiology laboratory. *Enferm. Infecc. Microbiol. Clin.* 39, 192–200. doi: 10.1016/J.EIMC.2020.02.027

Papagiannitsis, C. C., Študentová, V., Izdebski, R., Oikonomou, O., Pfeifer, Y., Petinaki, E., et al. (2015). Matrix-assisted laser desorption ionization-time of flight mass spectrometry Meropenem hydrolysis assay with NH_4HCO_3, a reliable tool for direct detection of Carbapenemase activity. *J. Clin. Microbiol.* 53, 1731–1735. doi: 10.1128/JCM.03094-14

Pierce, V. M., Simner, P. J., Lonsway, D. R., Roe-Carpenter, D. E., Johnson, J. K., Brasso, W. B., et al. (2017). Modified Carbapenem inactivation method for phenotypic detection of Carbapenemase production among *Enterobacteriaceae*. *J. Clin. Microbiol.* 55, 2321–2333. doi: 10.1128/JCM.00193-17

Potter, R. F., D'Souza, A. W., and Dantas, G. (2016). The rapid spread of carbapenem-resistant *Enterobacteriaceae*. *Drug Resist. Updat.* 29, 30–46. doi: 10.1016/j.drup.2016.09.002

Sakarikou, C., Ciotti, M., Dolfa, C., Angeletti, S., and Favalli, C. (2017). Rapid detection of carbapenemase-producing *Klebsiella pneumoniae* strains derived from blood cultures by matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS). *BMC Microbiol.* 17:54. doi: 10.1186/S12866-017-0952-3

Torres-Sangiao, E., Leal Rodriguez, C., and García-Riestra, C. (2021). Application and perspectives of MALDI–TOF mass spectrometry in clinical microbiology laboratories. *Microorganisms* 9:1539. doi: 10.3390/microorganisms9071539

van der Zwaluw, K., De Haan, A., Pluister, G. N., Bootsma, H. J., De Neeling, A. J., and Schouls, L. M. (2015). The Carbapenem inactivation method (CIM), a simple and low-cost alternative for the Carba NP test to assess phenotypic Carbapenemase activity in gram-negative rods. *PLoS One* 10:e0123690. doi: 10.1371/journal.pone.0123690

Yoon, E.-J., and Jeong, S. H. (2021). MALDI-TOF mass spectrometry technology as a tool for the rapid diagnosis of antimicrobial resistance in bacteria. *Antibiotics* 10:982. doi: 10.3390/antibiotics10080982

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 08 December 2022
DOI 10.3389/fmicb.2022.1057621

OPEN ACCESS

EDITED BY
Karsten Becker,
University Medicine Greifswald,
Germany

REVIEWED BY
Zachary R. Stromberg,
Pacific Northwest National Laboratory (DOE), United States
Timo Homeier-Bachmann,
Friedrich-Loeffler-Institute,
Germany

*CORRESPONDENCE
Matthew M. Hille
mhille@unl.edu

SPECIALTY SECTION
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

RECEIVED 29 September 2022
ACCEPTED 22 November 2022
PUBLISHED 08 December 2022

CITATION
Olson HG, Loy JD, Clawson ML, Wynn EL and Hille MM (2022) Genotype classification of *Moraxella bovis* using MALDI-TOF MS profiles. *Front. Microbiol.* 13:1057621. doi: 10.3389/fmicb.2022.1057621

Genotype classification of *Moraxella bovis* using MALDI-TOF MS profiles

Hannah G. Olson[1], John Dustin Loy[1], Michael L. Clawson[2], Emily L. Wynn[2] and Matthew M. Hille[1]*

[1]School of Veterinary Medicine and Biomedical Sciences, Institute for Agriculture and Natural Resources, University of Nebraska-Lincoln, Lincoln, NE, United States, [2]United States Department of Agricultural, Agricultural Research Service, United States Meat Animal Research Center, Clay Center, NE, United States

Moraxella bovis (*M. bovis*) is regarded as a causative agent of infectious bovine keratoconjunctivitis (IBK), the most common ocular disease of cattle. Recently, whole genome sequencing identified the presence of two distinct genotypes within *M. bovis* that differ in chromosome content, potential virulence factors, as well as prophage and plasmid profiles. It is unclear if the genotypes equally associate with IBK or if one is more likely to be isolated from IBK lesions. We utilized 39 strains of *M. bovis* that had previously undergone whole genome sequencing and genotype classification to determine the utility of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) to accurately genotype *M. bovis* strains. We successfully developed two biomarker models that accurately classified strains according to genotype with an overall accuracy of 85.8–100% depending upon the model and sample preparation method used. These models provide a practical tool to enable studies of genotype associations with disease, allow for epidemiological studies at the sub-species level, and can be used to enhance disease prevention strategies.

KEYWORDS

***Moraxella bovis*, MALDI-TOF MS, infectious bovine keratoconjunctivitis, bovine pinkeye, genotype**

Introduction

Infectious bovine keratoconjunctivitis (IBK) is the most common ocular disease of cattle (Brown et al., 1998; Kneipp, 2021). IBK clinically presents as a herd-level disease that is often seasonal and can occur with high morbidity (Kneipp, 2021). Clinical signs of IBK include corneal ulceration, lacrimation, conjunctivitis, blepharospasm and potential blindness in severe cases (Kneipp, 2021). *Moraxella bovis* (*M. bovis*) is regarded as the most strongly associated causal agent of IBK, as the disease can be reproduced experimentally in calves by inoculating the cornea with *M. bovis* (Rogers et al., 1987; Beard and Moore, 1994). Other infectious agents such as *Moraxella bovoculi* (*M. bovoculi*), *Mycoplasma bovoculi*, and bovine herpesvirus – type 1 are often recovered from lesions or are found associated with ocular disease, but thus far

experimental inoculation of calves with these agents has not produced clinical signs consistent with IBK (Rosenbusch and Ostle, 1986; George et al., 1988; Angelos et al., 2007; Angelos, 2010). Additional environmental factors such as face flies, ultraviolet light, dusty conditions, and tall grasses are also thought to play a role in IBK development (Hughes et al., 1965, 1968; Gerhardt et al., 1982; Hall, 1984; Kopecky et al., 1986; Smith, 2012; Maier et al., 2021a). A precise and current estimate of the economic cost of IBK is lacking, but previous studies have estimated the impact to be between $150–$226 million dollars in the United States alone (Killinger et al., 1977; Hansen, 2001; Dennis and Kneipp, 2021). The economic losses associated with IBK are due to the costs of treatment as well as decreased average daily gain in affected calves (Thrift and Overfield, 1974).

Prevention of IBK is often focused on vaccination and minimizing fly load (Sheedy et al., 2021). There are a number of fully licensed *M. bovis* vaccines, a single conditionally approved *M. bovoculi* vaccine, and autogenous vaccine formulations available from different manufacturers in the United States. Under experimental fields conditions, these vaccine formulations have all had mixed results in terms of preventing IBK regardless of the formulation, route of administration, or antigen makeup (Smith et al., 1990; Davidson and Stokka, 2003; O'Connor et al., 2011; Cullen et al., 2017; Hille et al., 2022).

In the United States, tetracycline and tulathromycin are the only antibiotics with label indications for IBK whereas florfenicol also has a label indication for IBK in Canada (Bio-Mycin 200 (oxytetracline) [package insert]. Duluth, GA; Boehringer Ingelheim Animal Health United States Inc. 2019, Draxxin (tulathromycin) [package insert] Kalamazoo, MI: Zoetis Inc. 2018, Nuflor (florfenicol) [package insert] Kirkland, Quebec, Canada, Merck Animal Health Intervet Canada Corp. 2019). The use of eye patches as an aid in treatment was recently shown to promote healing of corneal ulcers associated with IBK (Maier et al., 2021b).

A secreted repeats-in-toxin (RTX) exotoxin and a type IV pilus protein are the two main virulence factors possessed by *M. bovis* required for IBK development (Jayappa and Lehr, 1986; Clinkenbeard and Thiessen, 1991; Ruehl et al., 1993; Beard and Moore, 1994). Recently, two distinct genotypes of *M. bovis* were characterized that shared a core of 2,015 genes with an additional 121 genes specific to genotype 1 and 186 genes specific to genotype 2 (Wynn et al., 2022). The genotypes possess different sequence variants of RTX and different plasmid profiles. Specifically, only one genotype possessed plasmids containing filamentous hemagglutinin, a known virulence factor in other pathogens. These differences suggest the two genotypes may not be equally associated with IBK although this has not been proven.

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) is a commonly used method in diagnostic and research laboratories for the identification of bacteria (Seng et al., 2009; Clark et al., 2013; Sandalakis et al., 2017). Within *Moraxella* spp., MALDI-TOF MS has previously been used to accurately distinguish between *M. bovis* and *M. bovoculi* species, as well as distinguish genotypes within *M. bovoculi* (Robbins et al., 2018; Hille et al., 2020). Given the recent characterization of genotypes 1 and 2 of *M. bovis*, we hypothesized that MALDI-TOF MS may provide a timely and accurate method of genotype classification within this species as well. A rapid method to characterize strains of *M. bovis* according to genotype would allow for classification of a large number of disease associated strains to determine potential associations with disease in real-time as part of the bacterial identification process.

Materials and methods

Bacterial strains and culture conditions

Thirty seven of the 39 strains of *M. bovis* used for this study had previously undergone bacterial identification, whole genome sequencing and assembly into closed, circular chromosomes, and were classified according to genotype (Wynn et al., 2022). An additional two strains were genotyped from Illumina libraries using two methods: First, Illumina library fastq files were converted to fasta BLAST databases (Camacho et al., 2009) and previously identified genotype 1 and 2 specific genes (Wynn et al., 2022) were used as BLAST queries to find genotype specific gene sequence in the Illumina libraries. The strains consisted of ten genotype 1 and 29 genotype 2 strains. The strains primarily originated from diagnostic case submission samples from cattle with IBK that were submitted to the Nebraska Veterinary Diagnostic Center. The state of origin for each strain is shown in Table 1.

Frozen stocks of the strains were plated onto tryptic soy agar (TSA) with 5% sheep blood (Remel, Lenexa, KS) and incubated at 37°C for 24 h in 5% CO_2. The strains were then passed onto fresh blood agar plates incubated for another 24–48 h in the same conditions, and then pure colony growth was subjected to analysis by MALDI-TOF MS.

MALDI-TOF MS

MALDI TOF MS spectra was obtained for each of the strains per the manufacturer's recommendations using two methods, the smear method and the extraction method (Khot et al., 2012). To perform the smear method, a single colony was transferred onto the steel target plate using a wooden applicator and allowed to air dry before applying 1 µl of α-cyano-4-hydroxycinnamic acid (Bruker Daltonics, Billerica, MA). The wells were allowed to dry, and crystallization occurred. Analysis involved using MALDI Bioytyper system (Bruker Daltonik) in a positive linear mode with a mass range of 2–20 kDa m/z with laser frequency of 60 Hz and calibration using a Bacterial Test Standard (Bruker Daltonik). The first ion source had a voltage of 20,000 kV, and the second had a voltage of 18.10 kV with an additional lens voltage of 6.05 kV and a pulsed extraction time of 170 ns.

The extraction method involved using 2–3 colonies from solid media that were incubated for 24–48 h in 5% CO_2 at 37°C. After incubation, 300 µl of HPLC grade water and the colonies were

TABLE 1 State of origin and model groups assigned to the 39 strains used in this study.

Genotype	Model Group	Location
1	Generation	Nebraska (4)
		Florida (1)
1	Validation	Nebraska (2)
		Indiana (1)
		Saskatchewan, Canada (1)
		Wisconsin (1)
2	Generation	Kansas (4)
		Oregon (1)
		Minnesota (1)
		Nebraska (2)
		North Carolina (1)
		West Virginia (1)
		Montana (1)
		South Dakota (1)
		Iowa (1)
		Saskatchewan, Canada (1)
		Pennsylvania (1)
2	Validation	Nebraska (2)
		Kansas (1)
		Iowa (2)
		California (1)
		Texas (1)
		Illinois (1)
		Wisconsin (1)
		Florida (1)
		Illinois (1)
		Oregon (1)
		Oklahoma (1)
		West Virginia (1

The model group assignment was the same for developing both the GA and QC biomarker models.

vortexed until a homogenous mixture formed. Next, 900 µl of 100% ethanol was added and then centrifuged for 2 min at 16,000 xg. The supernatant was removed, and the pellet was allowed to air dry. Then, 25 µl of 70% formic acid and 25 µl of acetonitrile were combined with the pellet and centrifuged as mentioned above. Next, 1 µl of the supernatant was placed onto a well and allowed to air dry. The same 1 µl of matrix solution (α-cyano-4-hydroxycinnamic acid) was then added to each well before MALDI-TOF was performed. For each strain, eight wells were prepared using the extraction method, and each well was analyzed three times resulting in a total of 24 spectra. For the smear method, three wells were prepared and analyzed three times for a total of nine spectra. The spectra profiles were examined, and flat or inconsistent spectra were removed from the analysis.

Model generation and accuracy

Strains from each genotype were randomly assigned to either a biomarker model generation group or validation group (Table 1).

TABLE 2 Biomarker model characteristics and accuracy results for both the GA and QC biomarker models.

Model	GA	QC
Peaks used (m/z)	6,839	6,854
	6,854	
	7,301	
	8,769	
	9,103	
Extraction method		
Recognition Capability	100%	98.56%
Cross Validation	99.79%	98.21%
Classification genotype 1	100% (110/110)	100% (110/110)
Classification genotype 2	100% (307/307)	97.4% (299/307)
Smear method		
Classification genotype 1	81.3% (26/32)	100% (32/32)
Classification genotype 2	99.8% (122/123)	82.1% (101/123)

ClinProTools 3.0 software (Bruker Daltonik) was used to develop a biomarker model from the known genotypes within the model generation groups. Two classification algorithms were used to develop the models including genetic algorithm (GA) and quick classifier (QC). After the biomarker models were obtained, their accuracy was manually calculated according to the resulting classifications for each of the strains within the validation groups. The models were developed using spectra obtained via the extraction sample preparation method, and the accuracy of the models was assessed using spectra from both the extraction and smear sample preparation method. Any spectrum classified as "Null Spectrum" by the models was excluded from accuracy calculations. A two sample t-test assuming unequal variance was used to compare the classification accuracy between sample preparation methods and a two sample t-test assuming equal variance was used to compare classification accuracy between the models. The significance of genotype discrimination of individual peaks was determined using the "peak statistic" function within ClinProTools 3.0 after loading all spectra from each genotype.

A main spectrum profile (MSP) was created for each strain using 24 spectra from eight technical replicates using MBT Compass Explorer software (Bruker Daltonik). The MSP peak list function was used to determine the presence or absence of peaks included in the biomarker model for each strain. When present, the magnitude of each peak used in the biomarker model was recorded.

Results

The peaks included in the biomarker models for this study and the classification accuracies are summarized in Table 2. The GA biomarker model included five peaks and correctly classified 100% (110/110) and 100% (307/307) of validation group spectra obtained using the extraction method for genotype 1 and 2 strains, respectively. Therefore, the overall classification accuracy for extraction method spectra was 100% (417/417). When the smear method spectra were classified, the GA model correctly classified 81.3% (26/32) and 99.8%

TABLE 3 The presence and intensity of peak 6,854m/z used in the QC model for each of the validation group strains.

Peak (m/z)	6,854 (Range 6844.26–6864.35)	Average intensity (arbitrary units)
Extraction Method		
Genotype 1	1/5	3.01
Genotype 2	14/14	24.16
Smear Method		
Genotype 1	0/5	0.0
Genotype 2	14/14	38.38

(122/123) of the genotype 1 and 2 validation spectra, respectively, for an overall accuracy of 95.5% (146/155) which was significantly lower than the extraction method accuracy ($p = < 0.05$).

The QC biomarker model incorporated a single peak at 6854 m/z and correctly classified 100% (110/110) and 97.4% (299/307) of the genotype 1 and 2 extraction method validation spectra, respectively. The resulting overall classification accuracy for the extraction spectra was 98.1% (409/417). When classifying spectra obtained using the smear method, the QC model correctly classified 100% (32/32) and 82.1% (101/123) of genotype 1 and 2 spectra, respectively. The overall accuracy for smear spectra was therefore 85.8% (133/155) which was significantly lower than the extraction method accuracy ($p = < 0.05$). The accuracy of the GA model was statistically superior to the QC model using both the extraction method ($p = < 0.05$) and the smear method ($p = < 0.05$) of sample preparation. The highest weighted peak in the GA model was peak 6,854 m/z, which is the same peak used in the QC model. While the GA model showed superior accuracy, the QC model is still highly accurate and is appealing from a practical standpoint since it only uses a single peak. Highlighting the discriminatory power of this single peak would be especially useful if spectra were to be manually evaluated instead of using the ClinProTools 3.0 software. For this reason, we chose to focus on peak 6,854 m/z to examine the presence or absence, as well as the relative magnitude in the MSP profiles of all strains. Peak 6,854 m/z was highly significant between the two genotypes ($p = < 0.000001$) according to the Peak Statistic function within ClinProTools 3.0. This peak was present in all 14 genotype 2 validation MSPs with an average intensity of 24.16 arbitrary units (a.u.) and only present in 1/5 genotype 1 MSPs, and with a substantially lower intensity of only 3.01 a.u. (Table 3). Figure 1 displays the average spectra from each genotype at 6854 m/z as well as the distribution of each individual spectra.

Discussion

We successfully developed two MALDI-TOF MS biomarker models that accurately classified strains of *M. bovis* according to genotype. The GA model was significantly more accurate than the QC model using both the extraction and smear method. The accuracy of the smear method was significantly less than the extraction method for both models. However, the accuracy of the models using the smear method (95.5% for GA, 85.8% for QC) is likely sufficient given the ease of sample preparation compared to the extraction method. Regardless of the sample preparation method used, replication should be included to increase the discriminatory resolution of peak 6,854 m/z to account for individual profile variation of this peak, particularly when using the smear method. When the MSP of the strains were examined, determining the presence or absence of peak 6,854 m/z, in conjunction with the intensity, was sufficient to differentiate between the two genotypes. Therefore, manual observation of the MSP or individual spectrum profile of an unknown strain for a peak at 6854 m/z can allow for accurate genotype determination without the need for the biomarker model, if ClinProTools 3.0 software is not available.

One limitation for this study is that the collection of *M. bovis* strains with known genotypes is limited since they have only been recently described, and further study of the application of this model to strains more diverse in space and time is warranted. As there is not a standard number of strains or spectra required to generate MALDI-TOF biomarker models, we included 24 spectra from each strain in the model generation portion of this study to capture variability both between and within individual strain spectra. The classification accuracy of the models indicates consistent differences between the genotypes and indicates that these models are a valuable tool to genotype uncharacterized strains with a high degree of accuracy.

MALDI-TOF delivers several benefits over both whole genome sequencing and PCR. While the initial investment in MALDI-TOF capabilities is substantial, the reagents are fewer and costs associated with testing an individual strain is more cost effective and results are available more quickly. Additionally, MALDI-TOF has proven more accurate in identifying members of the genus *Moraxella* to the species level when compared to PCR (Robbins et al., 2018). Additionally, the raw data generated from the instrument as part of the species identification run can be used directly to identify genotypes, providing added value to existing data.

Additional work to determine the relative abundance of each genotype within healthy and IBK affected eyes using MALDI-TOF MS profiles will allow us to determine if either of the genotypes is more likely to be associated with disease. If it is determined that a specific genotype is more likely to be disease-associated, it may be beneficial to preferentially include such strains in future vaccine formulations, particularly for autogenous vaccines. Beyond disease association, the *M. bovis* genotyping models will help determine any geographic or seasonal differences in the abundance of the genotypes as well. Any differences determined in either geographic or seasonal distribution of the genotypes may provide another method for vaccine formulation customization. If both genotypes are represented equally among diseased eyes, the genotype classification will still prove useful in any efforts to decipher any potential differences in the mechanics of pathogenesis and/or the

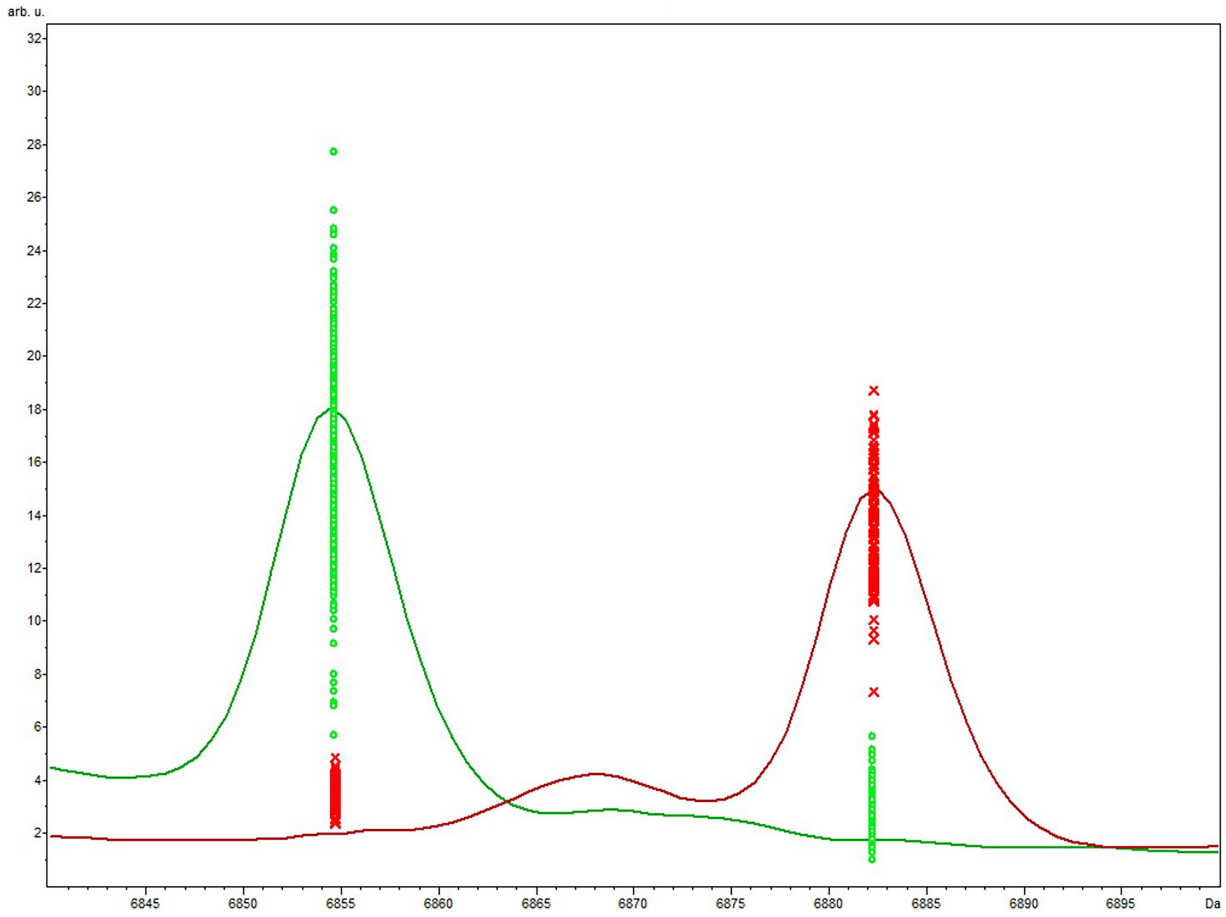


FIGURE 1
Average genotype 1 (red) vs. genotype 2 (green) spectra focused on the area of peak 6,854m/z included in the QC biomarker model. The differential expression and magnitude of peaks allows for differentiation of the respective genotypes.

utilization of certain virulence factors. The.XML files for both the GA and QC models developed in this study are available upon request by contacting the corresponding author.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Author contributions

MH and HO conceptualized the study, performed MALDI-TOF analysis, developed the biomarker models, wrote the first draft, and edited the manuscript. JL conceptualized the study and edited the manuscript. MC and EW genotyped the isolates and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

Funding for this study was provided by the Nebraska Experiment Station with funds from the Animal Health and Disease Research (section 1433) capacity funding program (accession 1017646) through the USDA National Institute of Food and Agriculture. The use of product and company names is necessary to accurately report the methods and results; however, the United States Department of Agriculture (USDA) neither guarantees nor warrants the standard of the products, and the use of names by the USDA implies no approval of the product to the exclusion of others that may also be suitable. The USDA is an equal opportunity provider and employer.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Angelos, J. A. (2010). *Moraxella bovoculi* and infectious bovine keratoconjunctivitis: cause or coincidence? *Vet. Clin. North Am. Food Anim. Pract.* 26, 73–78. doi: 10.1016/j.cvfa.2009.10.002

Angelos, J. A., Spinks, P. Q., Ball, L. M., and George, L. W. (2007). *Moraxella bovoculi* sp. nov., isolated from calves with infectious bovine keratoconjunctivitis. *Int. J. Syst. Evol. Microbiol.* 57, 789–795. doi: 10.1099/ijs.0.64333-0

Beard, M. K., and Moore, L. J. (1994). Reproduction of bovine keratoconjunctivitis with a purified haemolytic and cytotoxic fraction of *Moraxella bovis*. *Vet. Microbiol.* 42, 15–33.

Brown, M. H., Brightman, A. H., Fenwick, B. W., and Rider, M. A. (1998). Infectious bovine keratoconjunctivitis: a review. *J. Vet. Intern. Med.* 12, 259–266.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinform.* 10:421. doi: 10.1186/1471-2105-10-421

Clark, A. E., Kaleta, E. J., Arora, A., and Wolk, D. M. (2013). Matrix-assisted laser desorption ionization-time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. *Clin. Microbiol. Rev.* 26, 547–603. doi: 10.1128/CMR.00072-12

Clinkenbeard, K. D., and Thiessen, A. E. (1991). Mechanism of action of *Moraxella bovis* hemolysin. *Infect. Immun.* 59, 1148–1152.

Cullen, J. N., Engelken, T. J., Cooper, V., and O'Connor, A. M. (2017). Randomized blinded controlled trial to assess the association between a commercial vaccine against *Moraxella bovis* and the cumulative incidence of infectious bovine keratoconjunctivitis in beef calves. *J. Am. Vet. Med. Assoc.* 251, 345–351. doi: 10.2460/javma.251.3.345

Davidson, H. J., and Stokka, G. L. (2003). A field trial of autogenous *Moraxella bovis* bacterin administered through either subcutaneous or subconjunctival injection on the development of keratoconjunctivitis in a beef herd. *Can. Vet. J.* 44, 577–580.

Dennis, E. J., and Kneipp, M. (2021). A review of global prevalence and economic impacts of infectious bovine Keratoconjunctivitis. *Vet. Clin. North Am. Food Anim. Pract.* 37, 355–369. doi: 10.1016/j.cvfa.2021.03.010

George, L. W., Ardans, A., Mihalyi, J., and Guerra, M. R. (1988). Enhancement of infectious bovine keratoconjunctivitis by modified-live infectious bovine rhinotracheitis virus vaccine. *Am. J. Vet. Res.* 49, 1800–1806.

Gerhardt, R. R., Allen, J. W., Greene, W. H., and Smith, P. C. (1982). The role of face flies in an episode of infectious bovine keratoconjunctivitis. *J. Am. Vet. Med. Assoc.* 180, 156–159.

Hall, R. D. (1984). Relationship of the face fly (Diptera: Muscidae) to pinkeye in cattle: a review and synthesis of the relevant literature. *J. Med. Entomol.* 21, 361–365.

Hansen, R. (2001). NPew tools in the battle against pinkeye. Proceedings of Nevada Livestock Production Annual Update. Reno, NV.

Hille, M., Dickey, A., Robbins, K., Clawson, M. L., and Loy, J. D. (2020). Rapid differentiation of *Moraxella bovoculi* genotypes 1 and 2 using MALDI-TOF mass spectrometry profiles. *J. Microbiol. Methods* 173:105942. doi: 10.1016/j.mimet.2020.105942

Hille, M. M., Spangler, M. L., Clawson, M. L., Heath, K. D., Vu, H. L. X., Rogers, R. E. S., et al. (2022). A five year randomized controlled trial to assess the efficacy and antibody responses to a commercial and autogenous vaccine for the prevention of infectious bovine Keratoconjunctivitis. *Vaccine* 10:916. doi: 10.3390/vaccines10060916

Hughes, D. E., Pugh, G. W. Jr., and McDonald, T. J. (1965). Ultraviolet radiation and *Moraxella bovis* in the etiology of bovine infectious keratoconjunctivitis. *Am. J. Vet. Res.* 26, 1331–1338.

Hughes, D. E., Pugh, D. W., and McDonald, T. J. (1968). Experimental bovine infectious keratoconjunctivitis caused by sunlamp irradiation and *Moraxella bovis* infection: determination of optimal irradiation. *Am. J. Vet. Res.* 29, 821–827.

Jayappa, H. G., and Lehr, C. (1986). Pathogenicity and immunogenicity of piliated and nonpiliated phases of *Moraxella bovis* in calves. *Am. J. Vet. Res.* 47, 2217–2221.

Khot, P. D., Couturier, M. R., Wilson, A., Croft, A., and Fisher, M. A. (2012). Optimization of matrix-assisted laser desorption ionization-time of flight mass spectrometry analysis for bacterial identification. *J. Clin. Microbiol.* 50, 3845–3852. doi: 10.1128/JCM.00626-12

Killinger, A. H., Valentine, D., Mansfield, M. E., Ricketts, G. E., Cmarik, G. F., Neumann, A. H., et al. (1977). Economic impact of infectious bovine keratoconjunctivitis in beef calves. *Vet. Med. Small Anim. Clin.* 72, 618–620.

Kneipp, M. (2021). Defining and diagnosing infectious bovine Keratoconjunctivitis. *Vet. Clin. North Am. Food Anim. Pract.* 37, 237–252. doi: 10.1016/j.cvfa.2021.03.001

Kopecky, K. E., Pugh, G. W., and McDonald, T. J. (1986). Infectious bovine keratoconjunctivitis: contact transmission. *Am. J. Vet. Res.* 47, 622–624.

Maier, G. U., Davy, J. S., Forero, L. C., Bang, H., Clothier, K., and Angelos, J. A. (2021b). Effects of eye patches on corneal ulcer healing and weight gain in stocker steers on pasture: a randomized controlled trial. Translational. *Anim. Sci.* 5:txab162. doi: 10.1093/tas/txab162

Maier, G., Doan, B., and O'Connor, A. M. (2021a). The role of environmental factors in the epidemiology of infectious bovine Keratoconjunctivitis. *Vet. Clin. North Am. Food Anim. Pract.* 37, 309–320. doi: 10.1016/j.cvfa.2021.03.006

O'Connor, A. M., Brace, S., Gould, S., Dewell, R., and Engelken, T. (2011). A randomized clinical trial evaluating a farm-of-origin autogenous *Moraxella bovis* vaccine to control infectious bovine keratoconjunctivis (pinkeye) in beef cattle. *J. Vet. Intern. Med.* 25, 1447–1453. doi: 10.1111/j.1939-1676.2011.00803.x

Robbins, K., Dickey, A. M., Clawson, M. L., and Loy, J. D. (2018). Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry identification of *Moraxella bovoculi* and *Moraxella bovis* isolates from cattle. *J. Vet. Diagn. Investig.* 30, 739–742. doi: 10.1177/1040638718789725

Rogers, D. G., Cheville, N. F., and Pugh, G. W. (1987). Pathogenesis of corneal lesions caused by *Moraxella bovis* in gnotobiotic calves. *Vet. Pathol.* 24, 287–295.

Rosenbusch, R. F., and Ostle, A. G. (1986). *Mycoplasma bovoculi* infection increases ocular colonization by Moraxella ovis in calves. *Am. J. Vet. Res.* 47, 1214–1216.

Ruehl, W. W., Marrs, C. F., George, L., Banks, S. J., and Schoolnik, G. K. (1993). Infection rates, disease frequency, pilin gene rearrangement, and pilin expression in calves inoculated with *Moraxella bovis* pilin-specific isogenic variants. *Am. J. Vet. Res.* 54, 248–253.

Sandalakis, V., Goniotakis, I., Vranakis, I., Chochlakis, D., and Psaroulaki, A. (2017). Use of MALDI-TOF mass spectrometry in the battle against bacterial infectious diseases: recent achievements and future perspectives. *Expert Rev. Proteomics* 14, 253–267. doi: 10.1080/14789450.2017.1282825

Seng, P., Drancourt, M., Gouriet, F., La Scola, B., Fournier, P. E., Rolain, J. M., et al. (2009). Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin. Infect. Dis.* 49, 543–551. doi: 10.1086/600885

Sheedy, D. B., Samah, F. E., Garzon, A., Fausak, E., Van Noord, M., Angelos, J. A., et al. (2021). Non-antimicrobial approaches for the prevention or treatment of infectious bovine keratoconjunctivitis in cattle applicable to cow-calf operations: a scoping review. *Animal* 15:100245. doi: 10.1016/j.animal.2021.100245

Smith, S. (2012). Is Pinkeye a reason for clipping your pasture fields? Ohio BEEF Cattle Letter: The Ohio State University.

Smith, P. C., Blankenship, T., Hoover, T. R., Powe, T., and Wright, J. C. (1990). Effectiveness of two commercial infectious bovine keratoconjunctivitis vaccines. *Am. J. Vet. Res.* 51, 1147–1150.

Thrift, F. A., and Overfield, J. R. (1974). Impact of pinkeye (infectious bovine kerato-conjunctivitis) on weaning and postweaning performance of Hereford calves. *J. Anim. Sci.* 38, 1179–1184.

Wynn, E. L., Hille, M. M., Loy, J. D., Schuller, G., Kuhn, K. L., Dickey, A. M., et al. (2022). Whole genome sequencing of *Moraxella bovis* strains from North America reveals two genotypes with different genetic determinants. *BMC Microbiol.* 22, 258–270. doi: 10.1186/s12866-022-02670-3

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 22 February 2023
DOI 10.3389/fmicb.2023.1045289

OPEN ACCESS

EDITED BY
Karsten Becker,
University Medicine Greifswald,
Germany

REVIEWED BY
Xiaogang Xu,
Fudan University,
China
Antonella Lupetti,
University of Pisa,
Italy

*CORRESPONDENCE
Elena De Carolis
✉ elena.decarolis@policlinicogemelli.it

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

RECEIVED 15 September 2022
ACCEPTED 01 February 2023
PUBLISHED 22 February 2023

CITATION
Foglietta G, De Carolis E, Mattana G, Onori M, Agosta M, Niccolai C, Di Pilato V, Rossolini GM, Sanguinetti M, Perno CF and Bernaschi P (2023) "CORE" a new assay for rapid identification of *Klebsiella pneumoniae* COlistin REsistant strains by MALDI-TOF MS in positive-ion mode. *Front. Microbiol.* 14:1045289. doi: 10.3389/fmicb.2023.1045289

"CORE" a new assay for rapid identification of *Klebsiella pneumoniae* COlistin REsistant strains by MALDI-TOF MS in positive-ion mode

Gianluca Foglietta[1†], Elena De Carolis[2*†], Giordana Mattana[1], Manuela Onori[1], Marilena Agosta[1], Claudia Niccolai[3], Vincenzo Di Pilato[4], Gian Maria Rossolini[3,5], Maurizio Sanguinetti[2], Carlo Federico Perno[1] and Paola Bernaschi[1]

[1]Microbiology Unit and Diagnostic Immunology, Bambino Gesù Pediatric Hospital, IRCCS, Rome, Italy, [2]Microbiology Unit, Department of Laboratory Sciences and Infectious Diseases, Fondazione Policlinico Universitario A. Gemelli, IRCCS, Rome, Italy, [3]Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy, [4]Department of Surgical Sciences and Integrated Diagnostics (DISC), University of Genoa, Genoa, Italy, [5]Clinical Microbiology and Virology Unit, University Hospital Careggi, Florence, Italy

Due to the global spread of pan resistant organisms, colistin is actually considered as one of the last resort antibiotics against MDR and XDR bacterial infections. The emergence of colistin resistant strains has been observed worldwide in Gram-negative bacteria, such as *Enterobacteriaceae* and especially in *K. pneumoniae,* in association with increased morbidity and mortality. This landscape implies the exploration of novel assays able to target colistin resistant strains rapidly.

In this study, we developed and evaluated a new MALDI-TOF MS assay in positive-ion mode that allows quantitative or qualitative discrimination between colistin susceptible (18) or resistant (32) *K. pneumoniae* strains in 3 h by using the "Autof MS 1000" mass spectrometer. The proposed assay, if integrated in the diagnostic workflow, may be of help for the antimicrobial stewardship and the control of the spread of *K. pneumoniae* colistin resistant isolates in hospital settings.

KEYWORDS

colistin resistance detection, *Klebsiella pneumoniae*, positive-ion mode, MALDI-TOF MS, rapid assay

Introduction

In the last decades, the global spread of carbapenemase-producing *Enterobacterales* (CPE), primarily *Klebsiella pneumoniae* producing KPC-type carbapenemase (KPC-Kp), posed urgent threats on public health (Tzouvelekis et al., 2012; Murray et al., 2022), accounting for difficult-to-treat infections associated with high mortality rates (Cassini et al., 2019).

Owing to the significant burden of disease and limited treatment options, Centers for Disease Control and Prevention (CDC) (2019) and World Health Organization (WHO) (2017) ranked CPE as 'critical-priority' pathogens to which address the development of novel antimicrobial compounds.

Although several new antimicrobial agents active against CPE have been recently approved and marketed, including the novel β-lactam/β-lactamase inhibitor combinations, older molecules as colistin still hold a place in the antibiotic armamentarium as salvage therapy for patients infected with multi-drug resistant (MDR) or extensively drug-resistant (XDR) organisms (Jean et al., 2019).

Colistin is a positive charged polypeptide antibiotic, belonging to the polymyxin class, which targets the lipopolysaccharide (LPS) moiety on the outer membrane gram negatives, inducing a displacement of cations by electrostatic interaction and thus causing the disruption and loss of cell membrane integrity. To counteract these effects, bacteria have evolved multiple adaptative strategies including chromosomal mutations in the genes associated with the modification pathways of the lipid A, use of efflux pumps or capsule and the horizontal transfer of the plasmid-carried gene *mcr-1* (Granata and Petrosillo, 2017; Wang et al., 2018; Hamel et al., 2021).

To date, the main mechanism of resistance is the modification of the lipid A.

In *K. pneumoniae,* alterations involving the *mgrB* gene, along with mutations in *pmrAB* and *phoPQ* loci, have been reported as the most common mechanisms of colistin resistance accounting for modification of LPS (Pragasam et al., 2017; Giordano et al., 2018), by addition of phosphoethanolamine (PEtN) and 4-amino-4-deoxy-L-arabinose (L-Ara4N) residues to the phosphate groups of lipid A (Cannatelli et al., 2014; Poirel et al., 2015), while horizontal acquisition of mcr-like genes was observed less frequently.

Nowadays, broth microdilution (BMD) is considered as the reference method for antimicrobial susceptibility testing (AST) of colistin by Clinical and laboratory standard institute (CLSI) and European committee on antimicrobial susceptibility testing (EUCAST) join subcommittee (2016). However, these phenotypic methods do not match with the need of a timely detection of colistin resistance for patient's isolates as they imply turnaround times of 16–24 h. On the other hand, PCR-based molecular methods, although rapid, only provide information on the presence/absence of the genes involved in the resistance mechanisms, which not always correlates to the isolate's phenotype.

Very recently, a novel MALDI-TOF based method (i.e., the MALDIxin test) able to detect colistin resistance in about 15–30 min, thanks to a shift of the mass unit of the native lipid A present in the resistant bacterial strains, has been developed and validated for *E. coli*, *A. baumannii* (Dortet et al., 2018) and *K. pneumoniae* (Dortet et al., 2020).

Anyway, these assays require switching the MALDI-TOF MS machine to the negative ion mode, not the modality routinely used for bacterial identification. Besides the availability of such a technology (actually still rare in the majority of the microbiology labs), the assay needs pre and post switching additional calibrations.

Here, we aimed to develop and validate the "CORE" assay, a new MALDI-TOF-based test in positive ion mode for rapid prediction of colistin resistance in *K. pneumoniae*, relying on the detection of a mass spectrum profile with an identification score lower or ≥ 6 by using the Autof MS 1000 mass spectrometer (Autobio). The assay was evaluated both as a method for MIC prediction and as a screening tool for colistin resistance (quantitative and qualitative AST respectively) by comparison with BMD AST results.

Materials and methods

Strains collection

The study collection included 50 colistin-resistant ($n = 32$) and-susceptible ($n = 18$) *K. pneumoniae* isolates, cultured from blood ($n = 33$), urine ($n = 4$), rectal swabs ($n = 8$), tracheal broncho-aspirates ($n = 3$), cerebrospinal fluid ($n = 1$) and wound swab ($n = 1$). The isolates, collected within the 2016–2021 period, were part of national surveys ($n = 21$; Di Pilato et al., 2021), of previously published studies ($n = 9$; Cannatelli et al., 2014; Arena et al., 2016; Boncompagni et al., 2022) or of a local collection available at the University of Florence (Florence, Italy) ($n = 14$).

Determination of the colistin resistance mechanisms

Within the study collection, 32 out of 50 *K. pneumoniae* isolates were previously characterized by whole-genome sequencing (WGS), and genetic alterations associated with colistin resistance were formerly investigated (Cannatelli et al., 2014; Di Pilato et al., 2021; Boncompagni et al., 2022). The remaining colistin resistant isolates ($n = 18$) were screened for the presence of the most common *mcr* gene variants by Real-time PCR, including *mcr-1* and *mcr-2*, and additional *mcr* genes using specific primer/probes combinations (Coppi et al., 2018; Yang et al., 2018) (Table 1).

Colistin susceptibility testing

Colistin susceptibility testing of the clinical isolates for the CORE assay was performed using the BMD method by Liofilchem™ ComASP™ Colistin Test Panel (Liofilchem, Roseto degli Abruzzi, Te, Italy) containing the dried up antibiotic in 27-fold dilutions (0.25–16 mg/L) following the manufacturer instructions. MIC results were interpreted according to EUCAST Clinical Breakpoints (v12.0, 2022).

CORE assay

In order to obtain a fast and accurate qualitative and quantitative AST for colistin of the 50 *K. pneumoniae* strains included in the study, CORE assay and colistin BMD susceptibility testing were performed in parallel including three biological and two technical replicates for each tested isolate. To optimize the CORE assay inoculum concentration, the time of the test incubation and to develop a classifying algorithm, preliminary experiments were performed on 10 well-characterized *K. pneumoniae*, 5 colistin resistant (colR) and 5 colistin susceptible (colS) strains (reported in bold in Table 2).

The CORE assay protocol was performed according to the steps reported in Figure 1.

Briefly, all the *K. pneumoniae* strains were sub-cultured on MacConkey agar plates (bioMérieux) at 37° C for 24 h; a 0.5 McF suspension was made from grown colonies. Except for the inoculum size, ComASP (Compact Antimicrobial Susceptibility Panel) Colistin test was performed following the standard procedures.

TABLE 1 Sequence of primers and probes used to evaluate the presence of mcr-type genes.

Target genes	Primer name	Sequence 5′-3′	References
mcr-1 like	mcr-1-rt-fwd	ATCAGCCAAACCTATCCCATC	
	mcr-1-rt-rev	ACACAGGCTTTAGCACATAGC	Giani et al. (2013)
	mcr-1-rt-p	CY5-GACAATCTCGGCTTTGTGCTGACGATC-BHQ-3	
mcr-2 like	mcr-2-rt-fwd	AGCGATGGCGGTCTATCCTG	
	mcr-2-rt-rev	CAAAAAACGCCAAATTCATCAAGTC	Giani et al. (2013)
	mcr-2-rt-p	HEX-TGATGGGTGCTATGCTACTGATTGTCG-BHQ-1	
mcr-3 like	mcr-3-rt-fwd	CCAATCAAAATGAGGCGTTAGC	
	mcr-3-rt-rev	CACTATAAGTGATGCAAACATCG	*This study*
	mcr-3-rt-p	ROX-GGGCACGAGTTAGAATCCCTTTGAACC-BHQ-2	
mcr-4 like	mcr-4-rt-fwd	CAATTACCAATCTACTGCTGACTG	
	mcr-4-rt-rev	GTAACGCCTTAACTCACTGTTG	*This study*
	mcr-4-rt-p	FAM-CTGCTAATGTTCGTTGGCATTGGGATAG-BHQ-1	
mcr-5 like	mcr-5-rt-fwd	GCTGCCTGGATGAAATTCTGC	
	mcr-5-rt-rev	GTGTTCACCAAGGCTTCATGC	*This study*
	mcr-5-rt-p	CY5.5-CAGATGGGTGGTGTCGCAGGTTG-BBQ650	
mcr-6 like	mcr-6-rt-fwd	ACACAGCATAGTCCTTGGTAC	
	mcr-6-rt-rev	AACAGCACAGTAATCAATAGCATC	*This study*
	mcr-6-rt-p	FAM-CACCAATACTTATCCGATGGCACAAAAC-BHQ-1	
mcr-7 like	mcr-7-rt-fwd	TGGAGACCAACAACAGTGAG	
	mcr-7-rt-rev	CACGAACAGCAGCGAGAAGG	*This study*
	mcr-7-rt-p	HEX-TCGTGCTCTGGTTCCTGCTGAC-BHQ-1	
mcr-8 like	mcr-8-rt-fwd	CATCATACTTATCCGTTCCTTTTC	
	mcr-8-rt-rev	CCACAATTCAATTCTAAAAGCTCC	*This study*
	mcr-8-rt-p	ROX-GTACCAGCAATTATCCTGGCGTTGC-BHQ-2	
mcr-9 like	mcr-9-rt-fwd	ACGACTAAAGTGCCTTTCCAG	
	mcr-9-rt-rev	GATTCATATTCGAGAACATGCAC	*This study*
	mcr-9-rt-p	CY5-CTGGTAAAGGCATTGGTATCACGC-BHQ-3	

After 3 h of incubation at 37°C, an aliquot of 1 μL of the bacterial suspension from each well was spotted in duplicate on the MALDI-TOF target plate (96 wells steel target slide, Autobio Diagnostics) by direct deposition with the addition of 1 μL of Autobio sample pretreatment reagent lysate 1 (formic acid). After drying, the spot was overlaid with 1 μL of Cyano 4 Hydroxycinnamic Acid CHCA AUTOF MS matrix (Autobio Diagnostics, CO., LTD, China) solubilized in lysate 2 (acetonitrile) and buffer (trifluoroacetic acid) as recommended.

The spectra acquisition was performed on an Autof 1,000 MS (Autobio Diagnostics, Zhengzhou, China) mass spectrometer in positive linear mode, at a laser frequency of 60 Hz, in the mass range 2–20 kDa, using the "microbe" automatic acquisition mode with an overall 240 laser-shot acquired by 40 shot steps for each spot. The mass spectrometer was calibrated with Autobio calibrating agent consisting of nine calibrating proteins, according to manufacturer's instructions.

The protein spectra were analyzed using the Autof Acquirer version 1.0.55 software and the library v2.0.61 was used for the peaks matching. The identification results were interpreted according to the manufacturer criteria as following: Identification scores ≥9 were considered positive at the species level, scores between 6 and 9 as positive at the genus level, and scores <6 were defined as unreliable (no identification).

CORE assay data analysis

For the qualitative assay, only the spectra acquired by the plate well without the colistin agent and those acquired by the 2 mg/L plate well (ComASP® Colistin, Liofilchem, Te, Italy) were used as growth control and test breakpoint concentration (BP) respectively. An algorithm was developed to provide a rapid and accurate detection as colistin susceptible (colS) or resistant (colR) for the 50 *K. pneumoniae* strains. In particular, a test sample was classified as colR or colS on the following parameters: growth control spectra score ≥ 6 and 2 mg/L spectra score > 6 or 2 mg/L spectra score < 6, respectively.

On the contrary, for the quantitative assay all the spectra acquired in the range 0.25–16 mg/L were included in the analysis along with the growth control spectra. The MIC value was determined as the lowest drug dilution at which the spectra score was <6. The geometric mean (G MEAN) and the modal MIC calculated from the replicates was reported for all the samples (Table 2). In the case the 16 mg/L spectra

TABLE 2 Characteristics and results of the MALDI-TOF CORE assay on *Klebsiella pneumoniae* isolates included in the study (*n*=38).

Isolates[a]	Specimen	Mechanisms of colistin resistance	Acquired resistance to β-lactams	Phenotipic evaluation		MALDI-TOF "CORE" assay			
				MIC (mg/L)	Category[b]	[BP][c] Score	MIC		Category[d]
							GMean	Modal	
1300073725	Urine	–	–	> 16	R	> 6.0	> 16	>16	R
7001452909	Urine	PmrB (P95L)	OXA-1,CTX-M-15,KPC-3, SHV-28	> 16	R	> 6.0	> 16	> 16	R
7013697707	Tracheo aspirate	PmrA (A41T)	TEM 1D,CTX-M-15,KPC-3, SHV-28	> 16	R	> 6.0	> 16	> 16	R
7014248502	Blood	Δ*mgrB*	–	> 16	R	> 6.0	> 16	> 16	R
7014832004	Tracheo aspirate	–	–	4	R	> 6.0	7.13	4	R
7028826801	Tracheo aspirate	–	–	> 16	R	> 6.0	> 16	> 16	R
7029218302	Tracheo aspirate	–	–	16	R	> 6.0	> 16	> 16	R
7035291107	Tracheo aspirate	–	–	8	R	> 6.0	12.7	16	R
7042734501	Broncho aspirate	–	–	8	R	> 6.0	8.98	16	R
7042770204	Broncho aspirate	–	–	8	R	> 6.0	10.08	8	R
7048020202	Tracheo aspirate	–	–	16	R	> 6.0	> 16	> 16	R
7048595202	Tracheo aspirate	–	–	16	R	> 6.0	> 16	> 16	R
7050590801	Urine	Δ*mgrB*	KPC-3	16	R	> 6.0	16	16	R
7070100801	Blood	–	–	4	R	> 6.0	4.76	16	R
7066530202	Blood	Δ*mgrB*	KPC-3	> 16	R	> 6.0	> 16	> 16	R
7055043403	Broncho aspirate	–	–	16	R	> 6.0	> 16	> 16	R
7071493102	Blood	–	–	16	R	> 6.0	> 16	> 16	R
7078248702	Central venous catheter	Δ*mgrB*	TEM-1D, KPC-3	16	R	> 6.0	8.98	8	R
7079499102	Pharyngeal swab	–	–	16	R	> 6.0	12.7	16	R
KP348	Urine	Δ*mgrB*	KPC-3	16	R	> 6.0	> 16	> 16	R
KP350	Blood	–	–	16	R	> 6.0	12.7	16	R
KP338	Blood	–	–	8	R	> 6.0	10.37	16	R
KP369	Blood	–	–	16	R	> 6.0	16	16	R
B2	Blood	–	KPC-3	4	R	> 6.0	16	16	R
KP207-2	Blood	Δ*mgrB*	KPC-3	16	R	> 6.0	12.7	16	R
KP366	Blood	Δ*mgrB*	KPC-3	> 16	R	> 6.0	> 16	> 16	R
KP321	Blood	–	KPC-3	> 16	R	> 6.0	> 16	> 16	R
KP331	Blood	Δ*mgrB*	KPC-3	>16	R	> 6.0	> 16	> 16	R
KP292	Blood	Δ*mgrB*	KPC-3	> 16	R	> 6.0	> 16	> 16	R
KP295	Blood	Δ*mgrB*	KPC-3	> 16	R	> 6.0	> 16	> 16	R
KP260	Blood	Δ*mgrB*	KPC-3	> 16	R	> 6.0	> 16	> 16	R
KP266	Blood	Δ*mgrB*	KPC-3	> 16	R	> 6.0	> 16	> 16	R
KP377	Blood	–	KPC-2	0.25	S	< 6.0	0.56	0.5	S
KP290	Tracheo aspirate	–	–	0.5	S	< 6.0	0.65	0.5	S
KP291	Tracheo aspirate	–	–	0.5	S	< 6.0	0.63	0.5	S
KP317	Blood	–	KPC-2	1	S	< 6.0	1	1	S
KP330	Blood	Δ*mgrB*, pmrA (A217V), pmrB (G256R)	KPC-3,SHV1	< 0.25	S	< 6.0	0.79	0.5	S

(Continued)

10.3389/fmicb.2023.1045289

TABLE 2 (Continued)

Isolates[a]	Specimen	Mechanisms of colistin resistance	Acquired resistance to β-lactams	Phenotipic evaluation		MALDI-TOF "CORE" assay			
KP310	Blood	Δ*mgrB*	KPC-3	0.5	S	< 6.0	1	1	S
KP311	Blood	Δ*mgrB*	KPC-3	0.25	S	< 6.0	0.5	1	S
KP313	Blood	–	KPC-3	1	S	< 6.0	1.19	1	S
KP264	Blood	–	KPC-3	0.5	S	< 6.0	1.26	1	S
1300073725	Cerebral spinal fluid	–	KPC-3	0.25	S	< 6.0	0.35	0.25	S
7001452909	Blood	–	KPC-3	0.5	S	< 6.0	0.79	1	S
7013697707	Blood	–	KPC-3	< 0.25	S	< 6.0	1	1	S
7014248502	Blood	–	KPC-3	0.5	S	< 6.0	0.35	0.25	S
7014832004	Blood	–	KPC-3	0.5	S	< 6.0	1.19	1	S
7028826801	Blood	Δ*mgrB*	KPC-3	0.5	S	< 6.0	0.5	0.5	S
B1	Blood	–	KPC-3	< 0.25	S	< 6.0	< 0.25	< 0.25	S
KP207-1	Blood	–	KPC-2	< 0.25	S	< 6.0	< 0.25	< 0.25	S
KP293	Blood	–	KPC-3	0.5	S	< 6.0	0.5	0.5	S

Alterations of genes known to be involved in colistin resistance were reported for isolates previously characterized by WGS.

[a]In bold the selected colS and colR *K. pneumoniae* strains used for the preliminary experiments.

[b]Category were assigned according to EUCAST clinical breakpoints ver.12.02022 (S < 4, R ≥ 4 mg/L).

[c]Breakpoint concentration MALDI-TOF MS score.

[d]Category were assessed at the breakpoint concentration according to the proposed algorithm: MALDI-TOF MS score < 6.0 = S, ≥ 6.0 = *R*.

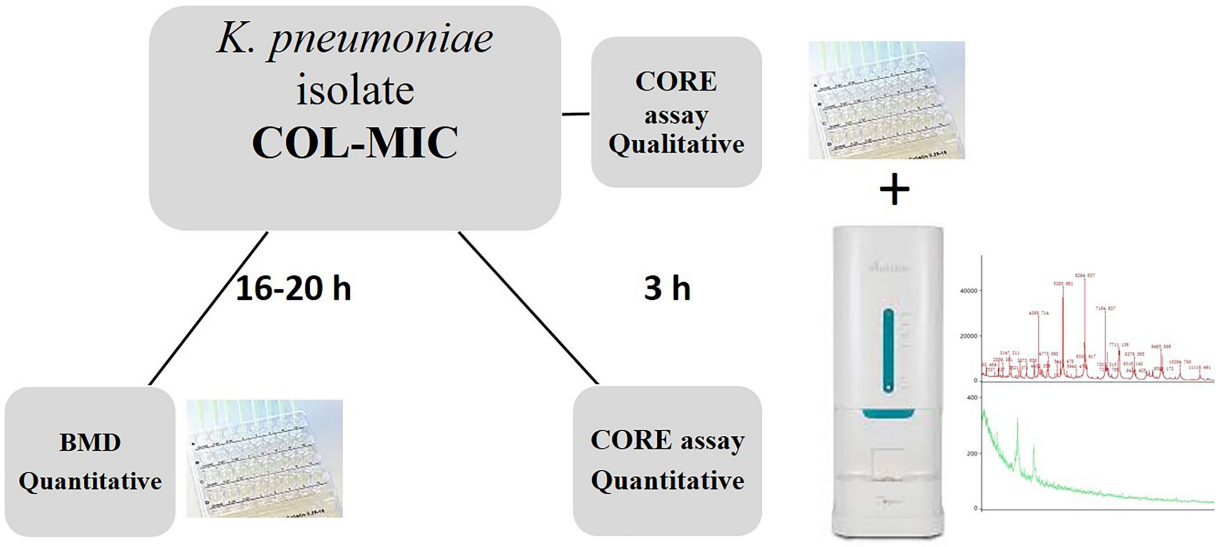


FIGURE 1
Laboratory workflow illustrating the CORE assay for the rapid detection of qualitative and quantitative colistin resistance in *K. pneumoniae* isolates.

score was >6, a MIC value >16 mg/L was reported and the test sample classified as colR. On the contrary, in the case the 0.25 mg/L spectra score was <6, a MIC value <0.25 mg/L was reported and the sample categorized as colS.

Thus, each strain was classified according to the above mentioned algorithms for the qualitative or quantitative CORE assay.

The results were compared with those obtained by the conventional BMD test, following interpretation with the EUCAST Clinical Breakpoints (v12.0, 2022) for colistin.

According to the modal MIC value, a MIC value agreement within ±1 dilution against BMD (essential agreement, EA) was considered acceptable (Clinical laboratory testing and in vitro diagnostic test systems ISO 20776-2:2021) for the new assay evaluation.

Divergence degree in distribution between BMD and CORE assay MIC values, potentially leading to resistance state misclassification, has been statistically evaluated with a non parametric Wilcoxon Test (R statistics stats library) performed in a paired manner and with a value of p continuity correction.

Results and discussion

The preliminary validation of the CORE assay carried out on 10 selected well-characterized strains of *K. pneumoniae* (i.e., including five colistine susceptible and five resistant strains, reported in bold in the Table 2), allow to correctly classifying them as colistin susceptible or resistant by the proposed algorithm.

The spectra of a colistin resistant and a colistin susceptible *K. pneumoniae* selected strains acquired by the Autof MS 1000 mass spectrometer using the CORE assay protocol are shown in Figure 2. The correspondent results obtained by the automatic matching of the acquired mass spectra against the Autobio library v2.0.61 for the growth control (no drug), BP and maximum concentration of colistin are reported at each profile side.

As exemplified, in the case of a susceptible *K. pneumoniae* isolate (Figure 2A), a matching score result <6.0 is obtained for all the tested replicates, both at the breakpoint (2 mg/L) and at the maximum concentration (16 mg/L), which takes into account the reduction or absence of the mass peaks in the respective MALDI-TOF MS profiles, in comparison with the growth control.

Conversely, for a resistant *K. pneumoniae* isolate (Figure 2B), the matching score results are above 6.0 both for the BP and maximum concentration, thus indicating the presence of a colistin resistant organism associated with the persistence of the mass spectra profiles at the given concentrations.

Thus, the CORE assay was applied on the 50 clinical isolates of *K. pneumoniae* included in the study. Genomic data revealed that multiple alterations in genes known to be involved in colistin resistance (i.e., *mgrB*, *pmrB*) were present in sequenced isolates ($n = 32$), regardless of the colistin resistance phenotype (Table 2), while no acquired *mcr*-like genes were detected in the whole isolate collection.

The CORE assay results for the qualitative or quantitative assay were compared with those obtained by the BMD quantitative test; the overall results are shown in Table 2. As reported all *K. pneumoniae*, 32 colR and 18 colS isolates, were correctly classified in 3 h as resistant or susceptible by the CORE assay, respectively. For what concerns the quantitative CORE assay, 30 out of 32 colR *K. pneumoniae* agreed against BMD MIC values within ±1 dilution according to the modal MIC results, whilst 7070100801 and B2 isolates obtained a MIC value 2 dilution higher (16 vs. 4 mg/L) using the quantitative CORE assay. Overall, an EA of 93.7% (30/32 *K. pneumoniae* resistant isolates) was reported following the 3 h incubation of the quantitative CORE assay.

Regarding the 18 colS *K. pneumoniae*, an EA of 83.3% was calculated; in particular, 15 out of 18 isolates resulted concordant against the BMD assay within ±1 dilution. However, KP338 MIC value was <0.25 mg/L instead of 0.5, KP377 obtained a MIC value 2 dilution higher and KP310 MIC value was <0.25 mg/L instead of 1 by comparison of quantitative CORE assay and BMD MIC results, respectively. Interestingly, all colistin susceptible isolates carrying genetic alterations previously associated with colistin resistance (i.e., Δ*mgrB*, $pmrA^{A217V}$, $pmrB^{G256R}$) were classified as colS by the "CORE" assay, a result consistent with the reference BMD (Table 2),

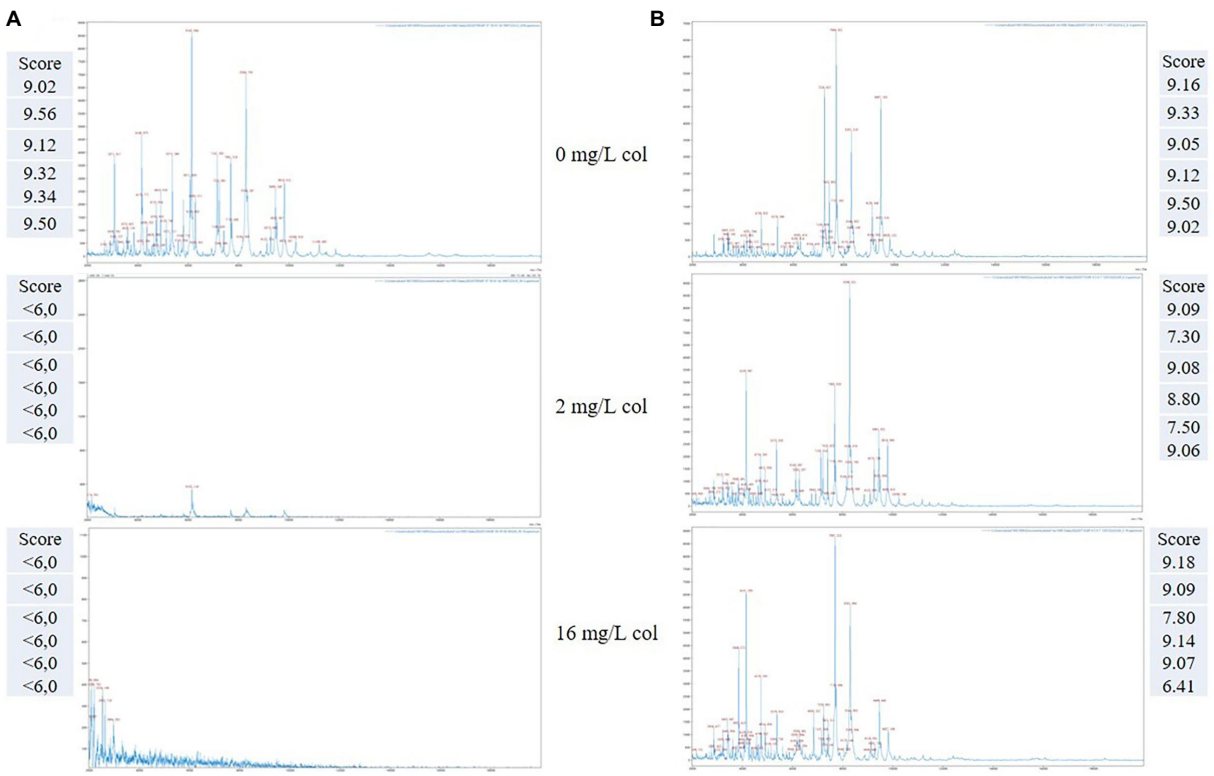


FIGURE 2
Representative MALDI-TOF mass spectra of two *K. pneumoniae* organisms detected as colistin susceptible **(A)** and colistin resistant **(B)**, respectively, by the CORE assay. Score value of the replicates spectra acquired at the selected drug concentrations are reported at the respective image side.

suggesting that colistin susceptibility could be more accurately predicted by the MALDI-TOF based approach than genetic data in these isolates.

The distribution divergence between BMD MIC values and CORE assay ones, quantitatively tested by mean of a paired Wilcoxon Test, estimated it as statistically significant since resulting value of p was 4.108e-05. This result indicates an overall agreement between mentioned approaches while considering MIC divergence degree.

In summary, following a three-hour samples incubation, by a simple algorithm for MALDI-TOF MS analysis (MALDI-TOF MS score < 6.0 = S or ≥ 6.0 = R), we obtained a total agreement between the qualitative CORE assay and the phenotypic method results. Moreover, an EA of 93.7 and 83.3% was achieved in the case of the quantitative CORE assay for the colR and colS isolates, respectively.

The MALDI-TOF MS based CORE assay in positive-ion mode that allows in 3 h of incubation the detection of colistin resistant or colistin susceptible *K. pneumoniae* isolates can provide rapid results to clinicians without the need to wait the 16–24 h necessary for the conventional BMD assays. Furthermore, the possibility to obtain colistin resistance detection using MALDI-TOF spectrometry instrument at the same polarity, the positive one, without switching to the negative ion mode and thus avoiding the calibration steps, makes the proposed algorithm suitable for a combination with the current routine MALDI-TOF MS identification workflow.

The simple algorithm here proposed for the CORE assay avoids statistical analysis based on the absence or presence of specific mass peaks and is independent from the mechanism of resistance, it relies on the growth of the microorganism and thus on the MALDI-TOF identification score matching value. Moreover, the test, can be suitable also for laboratories that cannot afford the costs of a new spectrometer equipped with the positive–negative ion mode switching modality as requested by the assays based on colistin resistance-related modifications to lipid A, thus consisting in a novelty in the landscape of polymixin resistance detection assays based on mass spectrometry.

One limitation of the qualitative CORE assay is that the time to result is higher with respect to the MALDI-TOF tests based on lipidomics. On the contrary, the quantitative CORE assay has the advantage to provide good results (EA of 93.7 and 83.3% for colR or colS *K. pneumoniae* isolates) earlier than conventional BMD methods (3 vs. 16–24 h).

Overall, the CORE assay, although studies are still needed to implement the number of samples tested, might be of extreme importance in the detection of colistin resistant isolates. The emergence of colistin resistant *K. pneumoniae* isolates can have a huge impact on the patient outcome representing a pressing health-care problem from both a management and economic point of view. The use of rapid and cost-effective new technologies applied to the resistance landscape, might offer the possibility to overcoming the non-appropriate use of colistin and at the same time can be of help in the struggle against the spread of antibiotic-resistance.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

EDC: conceived and planned the experiments. GF, GM, MO, and CN: carried out the experiments. MA: contributed to sample preparation. EDC, GF, VDP, and GM: contributed to the interpretation of the results. EDC: wrote the manuscript. VDP and CFP: provided critical feedback. GMR, MS CFP, and PB: supervised the project. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by EU funding within the MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project no. PE00000007, INF-ACT).

Acknowledgments

We wish to thank Valentino Costabile for his assistance in the statistical data analysis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AL declared a past co-authorship with the authors GMR, VDP to the handling editor at the time of review.

Publisher's note

References

Arena, F., Henrici De Angelis, L., Cannatelli, A., Di Pilato, V., Amorese, M., D'Andrea, M. M., et al. (2016). Colistin resistance caused by inactivation of the MgrB regulator is not associated with decreased virulence of sequence type 258 KPC Carbapenemase-producing Klebsiella pneumoniae. *Antimicrob. Agents Chemother.* 60, 2509–2512. doi: 10.1128/AAC.02981-15

Boncompagni, S. R., Micieli, M., Di Maggio, T., Aiezza, N., Antonelli, A., Giani, T., et al. (2022). Activity of fosfomycin/colistin combinations against planktonic and biofilm gram-negative pathogens. *J. Antimicrob. Chemother.* 77, 2199–2208. doi: 10.1093/jac/dkac142 Epub ahead of print

Cannatelli, A., Giani, T., D'Andrea, M. M., Di Pilato, V., Arena, F., Conte, V., et al. (2014). MgrB inactivation is a common mechanism of colistin resistance in KPC-producing Klebsiella pneumoniae of clinical origin. *Antimicrob. Agents Chemother.* 58, 5696–5703. doi: 10.1128/AAC.03110-14

Cassini, A., Högberg, L. D., Plachouras, D., Quattrocchi, A., Hoxha, A., Simonsen, G. S., et al. (2019). Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European economic area in 2015: a population-level modelling analysis. *Lancet Infect. Dis.* 19, 56–66. doi: 10.1016/S1473-3099(18)30605-4

Centers for Disease Control and Prevention (CDC). Antimicrobial resistance threats report (2019). Available from: https://www.cdc.gov/drugresistance/pdf/threats-report/2019-ar-threats-report-508. pdf (Accessed September 1, 2022)

Clinical and laboratory standard institute (CLSI) and European committee on antimicrobial susceptibility testing (EUCAST) join subcommittee. Recommendations for MIC determination of colistin (polymyxin E) as recommended by the joint CLSI-EUCAST Polymyxin Breakpoints Working Group (2016).

Clinical laboratory testing and in vitro diagnostic test systems — Susceptibility testing of infectious agents and evaluation of performance of antimicrobial susceptibility test devices — part 2: evaluation of performance of antimicrobial susceptibility test devices against reference broth micro-dilution. (2021). ISO 20776-2.

Coppi, M., Cannatelli, A., Antonelli, A., Baccani, I., Di Pilato, V., Sennati, S., et al. (2018). A simple phenotypic method for screening of MCR-1-mediated colistin resistance. *Clin. Microbiol. Infect.* 24:11. doi: 10.1016/j.cmi.2017.08.011

Di Pilato, V., Errico, G., Monaco, M., Giani, T., Del Grosso, M., Antonelli, A., et al. (2021). AR-ISS Laboratory study group on carbapenemase-producing Klebsiella pneumoniae. The changing epidemiology of carbapenemase-producing Klebsiella pneumoniae in Italy: toward polyclonal evolution with emergence of high-risk lineages. *J. Antimicrob. Chemother.* 76, 355–361. doi: 10.1093/jac/dkaa431, PMID: 33188415

Dortet, L., Brodam, A., Bernabeu, S., Glupczynski, Y., Bogaerts, P., Bonnin, R., et al. (2020). Optimization of the MALDIxin test for the rapid identification of colistin resistance in Klebsiella pneumoniae using MALDI-TOF MS. *J. Antimicrob. Chemother.* 75, 110–116. doi: 10.1093/jac/dkz405

Dortet, L., Potron, A., Bonnin, R. A., Plesiat, P., Naas, T., Filloux, A., et al. (2018). Rapid detection of colistin resistance in Acinetobacter baumannii using MALDI-TOF-based lipidomics on intact bacteria. *Sci. Rep.* 8:16910. doi: 10.1038/s41598-018-35041-y

Giani, T., Antonelli, A., Caltagirone, M., Mauri, C., Nicchi, J., Arena, F., et al. (2013). Evolving beta-lactamase epidemiology in Enterobacteriaceae from Italian nationwide surveillance, October: KPC-carbapenemase spreading among outpatients. *Euro Surveill.* 22:30583. doi: 10.2807/1560-7917.ES.2017.22.31.30583

Giordano, C., Barnini, S., Tsioutis, C., Chlebowicz, M. A., Scoulica, E. V., Gikas, A., et al. (2018). Expansion of KPC-producing Klebsiella pneumoniae with various mgrB mutations giving rise to colistin resistance: the role of IS L3 on plasmids. *Int. J. Antimicrob. Agents* 51, 260–265. doi: 10.1016/j.ijantimicag.2017.10.011

Granata, G., and Petrosillo, N. (2017). Resistance to colistin in Klebsiella pneumoniae: a 4.0 strain? *Infect Dis Rep* 9:7104. doi: 10.4081/idr.2017.7104

Hamel, M., Rolain, J. M., and Baron, S. A. (2021). The history of Colistin resistance mechanisms in bacteria. *Prog. Challeng. Microorgan.* 9:442. doi: 10.3390/microorganisms9020442

Jean, S. S., Gould, I. M., Lee, W. S., and Hsueh, P. R. (2019). International Society of Antimicrobial Chemotherapy (ISAC). New drugs for multidrug-resistant gram-negative organisms: time for stewardship. *Drugs* 79, 705–714. doi: 10.1007/s40265-019-01112-1

Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., et al. (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* 399, 629–655. doi: 10.1016/S0140-6736(21)02724-0

Poirel, L., Jayol, A., Bontron, S., Villegas, M. V., Ozdamar, M., Turkoglu, S., et al. (2015). The mgrB gene as a key target for acquired resistance to colistin in Klebsiella pneumoniae. *J. Antimicrob. Chemother.* 70, 75–80. doi: 10.1093/jac/dku323

Pragasam, A. K., Shankar, C., Veeraraghavan, B., Biswas, I., Nabarro, L. E., Inbanathan, F. Y., et al. (2017). Molecular mechanisms of Colistin resistance in Klebsiella pneumoniae causing bacteremia from India-a first report. *Front. Microbiol.* 7:2135. doi: 10.3389/fmicb.2016.02135

Tzouvelekis, L. S., Markogiannakis, A., Psichogiou, M., Tassios, P. T., and Daikos, G. L. (2012). Carbapenemases in Klebsiella pneumoniae and other Enterobacteriaceae: an evolving crisis of global dimensions. *Clin. Microbiol. Rev.* 25, 682–707. doi: 10.1128/CMR.05035-11

Wang, X., Wang, Y., Zhou, Y., Li, J., Yin, W., Wang, S., et al. (2018). Emergence of a novel mobile colistin resistance gene, mcr-8 in NDM-producing Klebsiella pneumoniae. *Emerg. Microb. Infect.* 7:122. doi: 10.1038/s41426-018-0124-z

World Health Organization (WHO). *Global Priority List of Antibiotic-resistant Bacteria to Guide Research, Discovery, and Development of New Antibiotics* (2017). Geneva: World Health Organization (WHO).

Yang, Y. Q., Li, Y. X., Lei, C. W., Zhang, A. Y., and Wang, H. N. (2018). Novel plasmid-mediated colistin resistance gene mcr-7.1 in Klebsiella pneumoniae. *J. Antimicrob. Chemother.* 73, 1791–1795. doi: 10.1093/jac/dky111

Frontiers in Microbiology

Explores the habitable world and the potential of microbial life

The largest and most cited microbiology journal which advances our understanding of the role microbes play in addressing global challenges such as healthcare, food security, and climate change.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact