MDPI

Article

Differentiation of *Bacillus cereus* and *Bacillus thuringiensis* Using Genome-Guided MALDI-TOF MS Based on Variations in Ribosomal Proteins

Minling Chen [1,2,†]**, Xianhu Wei** [2,†]**, Junhui Zhang** [3]**, Huan Zhou** [3]**, Nuo Chen** [3]**, Juan Wang** [2]**, Ying Feng** [2]**, Shubo Yu** [2]**, Jumei Zhang** [2]**, Shi Wu** [2]**, Qinghua Ye** [2]**, Rui Pang** [2]**, Yu Ding** [3,*] **and Qingping Wu** [2,*]

[1] School of Bioscience and Bioengineering, South China University of Technology, Guangzhou 510006, China; cmling0101@163.com

[2] Guangdong Provincial Key Laboratory of Microbial Safety and Health, State Key Laboratory of Applied Microbiology Southern China, Institute of Microbiology, Guangdong Academy of Sciences, Guangzhou 510070, China; wxhu7508@163.com (X.W.); wangjuan@scau.edu.cn (J.W.); fengying717@163.com (Y.F.); yushubo5663860@163.com (S.Y.); zhangjm@gdim.cn (J.Z.); wushiloveyou@126.com (S.W.); yeqinghua2002@163.com (Q.Y.); pr839@163.com (R.P.)

[3] Department of Food Science and Technology, Jinan University, Guangzhou 510632, China; conanjunhui@163.com (J.Z.); zh71739601@163.com (H.Z.); chennuo7676@163.com (N.C.)

[*] Correspondence: dingyu@gdim.cn (Y.D.); wuqp203@163.com (Q.W.)

[†] These authors contributed equally to this work.

Abstract: *Bacillus cereus* and *B. thuringiensis* are closely related species that are relevant to foodborne diseases and biopesticides, respectively. Unambiguous differentiation of these two species is crucial for bacterial taxonomy. As genome analysis offers an objective but time-consuming classification of *B. cereus* and *B. thuringiensis*, in the present study, matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) was used to accelerate this process. By combining in silico genome analysis and MALDI-TOF MS measurements, four species-specific peaks of *B. cereus* and *B. thuringiensis* were screened and identified. The species-specific peaks of *B. cereus* were m/z 3211, 6427, 9188, and 9214, and the species-specific peaks of *B. thuringiensis* were m/z 3218, 6441, 9160, and 9229. All the above peaks represent ribosomal proteins, which are conserved and consistent with the phylogenetic relationship between *B. cereus* and *B. thuringiensis*. The specificity of the peaks was robustly verified using common foodborne pathogens. Thus, we concluded that genome-guided MALDI-TOF MS allows high-throughput differentiation of *B. cereus* and *B. thuringiensis* and provides a framework for differentiating other closely related species.

Keywords: *Bacillus cereus*; *Bacillus thuringiensis*; genome-guided MALDI-TOF MS; ribosomal protein; high-throughput differentiation

1. Introduction

The *Bacillus cereus* group comprises many genetically highly related species, including *B. cereus* and *B. thuringiensis* [1]. These species are unavoidable in food production and processing because of their capability to form resistant spores as well as their ubiquitous distribution in the environment [2]. *B. cereus* is an opportunistic foodborne pathogen that can cause emetic and diarrheal symptoms owing to the production of cereulide and enterotoxins, respectively [3]. It has been estimated that *B. cereus* causes 63,400 foodborne illness cases in the United States every year [4]. From 2006 to 2016, *B. cereus* ranked third in the prevalence of foodborne pathogens in China [5]. *B. thuringiensis* is widely used as a biopesticide because of its ability to produce insecticidal proteins, including crystal (Cry) and cytolytic (Cyt) δ-endotoxins [6,7]. Recent studies have shown that *B. thuringiensis* can also be used to promote plant growth [8], produce metal nanoparticles [9], and biodegrade environmental pollutants [10]. Although *B. cereus* and *B. thuringiensis* have different

pathogenicity and applications, they are genetically closely related, and there is still no available ISO to reliably distinguish between them [1]. Accurate taxonomy of *B. cereus* and *B. thuringiensis* is of fundamental importance for phylogenetic taxonomy research, public health, food industry safety, biopesticide market, and biotechnology development.

Great efforts have been made to differentiate *B. cereus* and *B. thuringiensis*, however, the conventional method for differentiation between *B. cereus* and *B. thuringiensis* based on insecticidal crystal proteins has certain limitations, as the genome data show that insecticidal crystal protein genes are not unique to *B. thuringiensis* [11–13]. General molecular typing also cannot achieve reliable differentiation, as 16S rRNA gene sequences of *B. cereus* and *B. thuringiensis* share more than 99% similarity [14], and pulsed-field gel electrophoresis also cannot effectively differentiate between *B. cereus* and *B. thuringiensis* [15]. Several molecular targets were screened to identify strains within the *B. cereus* group [16–18], and among them, the *xre* gene showed specificity for detecting *B. thuringiensis* [16,18].

In the post-genomic era, whole genome sequencing has been increasingly used for bacterial taxonomy [19]. With the advantage of being reliable, reproducible, and objective [20], genome-based taxonomy offers phylogenetically consistent resolution for classifying bacterial species [21]. The Genome Taxonomy Database (GTDB) was established based on the genome phylogeny of 120 ubiquitous single-copy proteins and provided a standardized taxonomy for bacterial species, including *B. cereus* and *B. thuringiensis* [21]. The genome blast distance phylogeny approach was used to calculate intergenomic distances [22] and clearly separated *B. cereus* and *B. thuringiensis* [23]. In the same study, the phylogenetic tree of the housekeeping gene *pycA* (encoding pyruvate carboxylase) was found to resolve the confusion between *B. cereus* and *B. thuringiensis* [23].

Although genome-based taxonomy assists classification with high resolution, it is still time-consuming, expensive, and labor-intensive, and is not suitable for routine practice. Matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) has been extensively applied in clinical diagnosis [24], food safety control [25,26], and environmental monitoring [27], offering reliable, rapid, and cost-effective microbial identification. In recent studies, MALDI-TOF MS showed sufficient resolution to differentiate between five predominant serovars of the non-typhoidal *Salmonella* [28], O157, O26, and O111 serotypes of *Escherichia coli* [29], different species of the genus *Listeria* [30], and lineages of methicillin-resistant *Staphylococcus aureus* [31]. Meanwhile, MALDI-TOF MS had also been applied to differentiate strains within *B. cereus* group [32,33]. Using MALDI-TOF MS, it is possible to detect ribosomal proteins [34], which are ideal targets for phylogenetic analysis. Proteotyping based on ribosomal proteins provided congruent resolution in comparison to molecular typing methods such as multilocus sequence typing (MLST) [35,36]. By using ribosomal multilocus sequence typing (rMLST) based on 53 ribosomal protein genes, the taxonomic status of diverse species, including *B. cereus* and *B. thuringiensis*, was unambiguously classified [37]. In the present study, as ribosomal proteins are dependable biomarkers in MALDI-TOF MS analysis for species identification, we first screened for robust variations in ribosomal proteins between *B. cereus* and *B. thuringiensis* based on genome analysis and then correlated them with species-specific features in MALDI-TOF MS profiling. The aim of this study was to develop a high-throughput MALDI-TOF MS-based tool for differentiating between *B. cereus* and *B. thuringiensis*.

2. Materials and Methods

2.1. Bacterial Strains

Two reference strains (*B. cereus* ATCC 14,579 and *B. thuringiensis* ATCC 10792) and 79 strains of *Bacillus* species isolated in our laboratory [38–41] were selected to screen for potential biomarkers. Detailed information on the strains are provided in Table S1. Thirteen common foodborne pathogens were used to verify the specificity of the biomarkers.

2.2. Genomic Data Mining

The genome sequences of 106 *B. cereus* and 175 *B. thuringiensis* strains (Table S2) were downloaded from the National Center for Biotechnology Information (NCBI) database to mine potential biomarkers. The definite classification of selected strains was determined using the GTDB (https://gtdb.ecogenomic.org/, accessed on 1 January 2021) and rMLST databases (https://pubmlst.org/rmlst, accessed on 1 January 2021). The genome sequences were annotated using the Prokka software [42]. The output GFF files were used to perform genome alignment using Roary [43] with an 80% sequence identity cut-off. Aligned sequences of ribosomal proteins were obtained using TBtools [44]. Theoretical molecular weights of ribosomal proteins were calculated from translated amino acid sequences using the web-based compute pI/Mw tool (https://web.expasy.org/compute_pi/, accessed on 15 January 2021) and protein molecular weight tool (http://www.detaibio.com/sms2/protein_mw.html, accessed on 15 January 2021). When the second residue in the amino acid chain was Gly, Ala, Pro, Ser, Thr, Val, or Cys, molecular weights were calculated by considering the N-terminal methionine loss.

2.3. Species Identification Based on Phylogenetic Tree of pycA

Genomic DNA was extracted using an ultrasonic-assisted method. Briefly, the bacterial colony suspended in 50 μL of ddH$_2$O was treated with an ultrasonic bath operating at 40 kHz for 5 min. The suspension was centrifuged at 15,000× *g* for 2 min. Then, 1 μL of supernatant was used for the following PCR reaction, the reaction mixture contained 12.5 μL of Premix Taq™ (TaKaRa Taq™ Version 2.0 plus dye, Guangzhou, China), 1 μL of each primers (10 μM), 1 μL of DNA template, and 9.5. μL of sterile water. PCR amplification of the *pycA* gene was carried out using the following primers: forward primer, 5′-GTGAAAGCAAGAACACAAGC-3′, and reverse primer, 5′-ATAGTTTTTGTATCCAACTGCG-3′. PCR reactions were conducted as follows: one cycle of initial denaturation at 98 °C for 3 min, followed by 35 cycles of 98 °C for 10 s, 55 °C for 30 s, and 72 °C for 100 s, and a final extension at 72 °C for 10 min. The PCR products were purified and sequenced by Tianyi Huiyuan Bioscience & Technology Inc. (Guangzhou, China).

Identification of *B. cereus* and *B. thuringiensis* was performed as described previously [45]. Briefly, the obtained sequences were aligned with the *pycA* genes of *B. cereus* group reference strains in Liu's study [23], after which a phylogenetic analysis using the maximum likelihood algorithm was performed with the MEGA X software [46]. Visualization of the phylogenetic tree was performed using iTOL [47].

2.4. MALDI-TOF MS Analysis

For MALDI-TOF MS analysis, *Campylobacter jejuni* was inoculated in Skirrow agar plates (Guangdong Huankai Co., Ltd., Guangzhou, China), *Vibrio parahaemolyticus* was inoculated in the Luria-Bertani (LB) agar (Guangdong Huankai Co., Ltd.,) with 3% NaCl, and the other strains were inoculated in the LB agar. All strains were incubated at 37 °C for 16 h. Bacterial colonies were smeared directly onto a 96-well MALDI target plate (Bruker Daltonics, Bremen, Germany) using sterile toothpicks and overlaid with 1 μL of 70% formic acid. After air-drying, 1 μL of the matrix solution containing 10 mg/mL of α-cyano-4-hydroxy-cinnamic acid (HCCA) in 50% (*v*/*v*) acetonitrile with 2.5% (*v*/*v*) trifluoroacetic acid was applied and allowed to dry. The detection was carried out on three different days in quadruplicates for each strain.

The mass spectra were acquired using a Microflex LT/SH smart MALDI-TOF MS (Bruker Daltonics, Bremen, Germany) equipped with a 200 Hz smartbeam solid-state laser and operated in positive linear mode. Mass spectra were automatically recorded within a mass range of 2–20 kDa with a total of 240 laser shots. A bacterial test standard (Bruker Daltonics, Bremen, Germany) was used for external spectral calibration before every experiment.

2.5. Data Processing and Identification of Biomarker Proteins

Raw mass spectra were converted to mzML files using the R package MALDIquant [48]. The output mzML files were imported into the Mass-Up software (Mass-Up, Vigo, Spain) for data processing [49]. Mass-Up provided intensity transformation (square root method), spectral smoothing (Savitzky-Golay method), baseline correction (TopHat method), standardization (Total Ion Current method), and peak detection (MassSpecWavelet method; SNR = 6, PeakScaleRange = 2, amp.Th = 0.0001). Output peak lists from the replicates of each strain were used to construct intra- and inter-sample matching using the forward method (tolerance = 300 ppm, reference type = AVG). Species-specific peaks were screened using the biomarker discovery module. Biomarker identification was performed by comparing the species-specific m/z against the theoretical m/z of differential ribosomal proteins in the above genome analysis. The identification result was further validated by searching against proteins below 10,000 Da from genome analysis and proteins with corresponding molecular weights from the UniProt database. All spectra were visualized using FlexAnalysis (v3.4, Bruker Daltonics, Bremen, Germany) after smoothing and baseline subtraction.

3. Results

3.1. Genome Analysis

In total, 106 strains of *B. cereus* and 175 strains of *B. thuringiensis* (Table S2) that possessed consistent classification in NCBI, GTDB, and rMLST were selected for biomarker mining. Among the theoretical molecular weights of 53 ribosomal proteins (Table S3), 13 of these (L30, L31 type B, S20, S6, S10, L18, S13, L15, L13, S7, L6, L5, and L3) showed mass variations between *B. cereus* and *B. thuringiensis*, with a sensitivity greater than 98% (Table 1). Eleven ribosomal proteins with molecular weights below 20,000 Da could serve as potential biomarkers for MALDI-TOF MS measurements.

Table 1. Theoretical molecular weight and sensitivity of ribosomal proteins with variations between *Bacillus cereus* and *Bacillus thuringiensis* from the genome analysis.

Gene	Annotation	Post-Translational Modification	Theoretical Molecular Weight (Da)		Sensitivity (%)	
			B. cereus	*B. thuringiensis*	*B. cereus*	*B. thuringiensis*
rpmD	50S ribosomal protein L30	methionine removed	6424.60	6438.62	100 (106/106)	100 (175/175)
rpmE2	50S ribosomal protein L31 type B	-	9184.33	9157.30	98.11 (104/106)	100 (175/175)
rpsT	30S ribosomal protein S20	methionine removed	9210.58	9226.58	100 (106/106)	99.43 (174/175)
rpsF	30S ribosomal protein S6	-	11,298.99	11,284.92	100 (106/106)	98.86 (173/175)
rpsJ	30S ribosomal protein S10	methionine removed	11,552.45	11,566.48	100 (106/106)	100 (175/175)
rplR	50S ribosomal protein L18	-	13,105.97	13,093.92	100 (106/106)	100 (175/175)
rpsM	30S ribosomal protein S13	methionine removed	13,687.83	13,661.79	100 (106/106)	100 (175/175)
rplO	50S ribosomal protein L15	-	15,477.74	15,521.79	100 (106/106)	100 (175/175)
rplM	50S ribosomal protein L13	-	16,427.96	16,457.98	100 (106/106)	100 (175/175)
rpsG	30S ribosomal protein S7	methionine removed	17,779.61	17,914.77	99.06 (105/106)	100 (175/175)
rplF	50S ribosomal protein L6	methionine removed	19,384.34	19,341.32	99.06 (105/106)	100 (175/175)
rplE	50S ribosomal protein L5	-	20,165.48	20,136.44	100 (106/106)	100 (175/175)
rplC	50S ribosomal protein L3	methionine removed	22,560.99	22,544.99	99.06 (105/106)	99.43 (174/175)

3.2. Identification of B. cereus and B. thuringiensis Based on the pycA Gene

As the phylogenetic analysis of the *pycA* gene was able to rapidly differentiate between *B. cereus* and *B. thuringiensis* [23], we used this gene as a criterion to select the correct strains for further experiments. The *pycA* amplicon sequences of *B. cereus*, *B. thuringiensis*, and reference strains were listed in Table S1. A phylogenetic tree was constructed based on

the *pycA* gene (Figure 1). *B. subtilis* ATCC 6051 was used to root the tree, and the strains within the *B. cereus* group in Liu's study were used as reference strains. Different strains of *B. cereus* or *B. thuringiensis* were clearly divided into two branches. The *pycA* gene sequence-based phylogenetic analysis assigned the isolated strains to definite *B. cereus* and *B. thuringiensis*, which were used for further MALDI-TOF MS measurements to discover potential biomarkers.
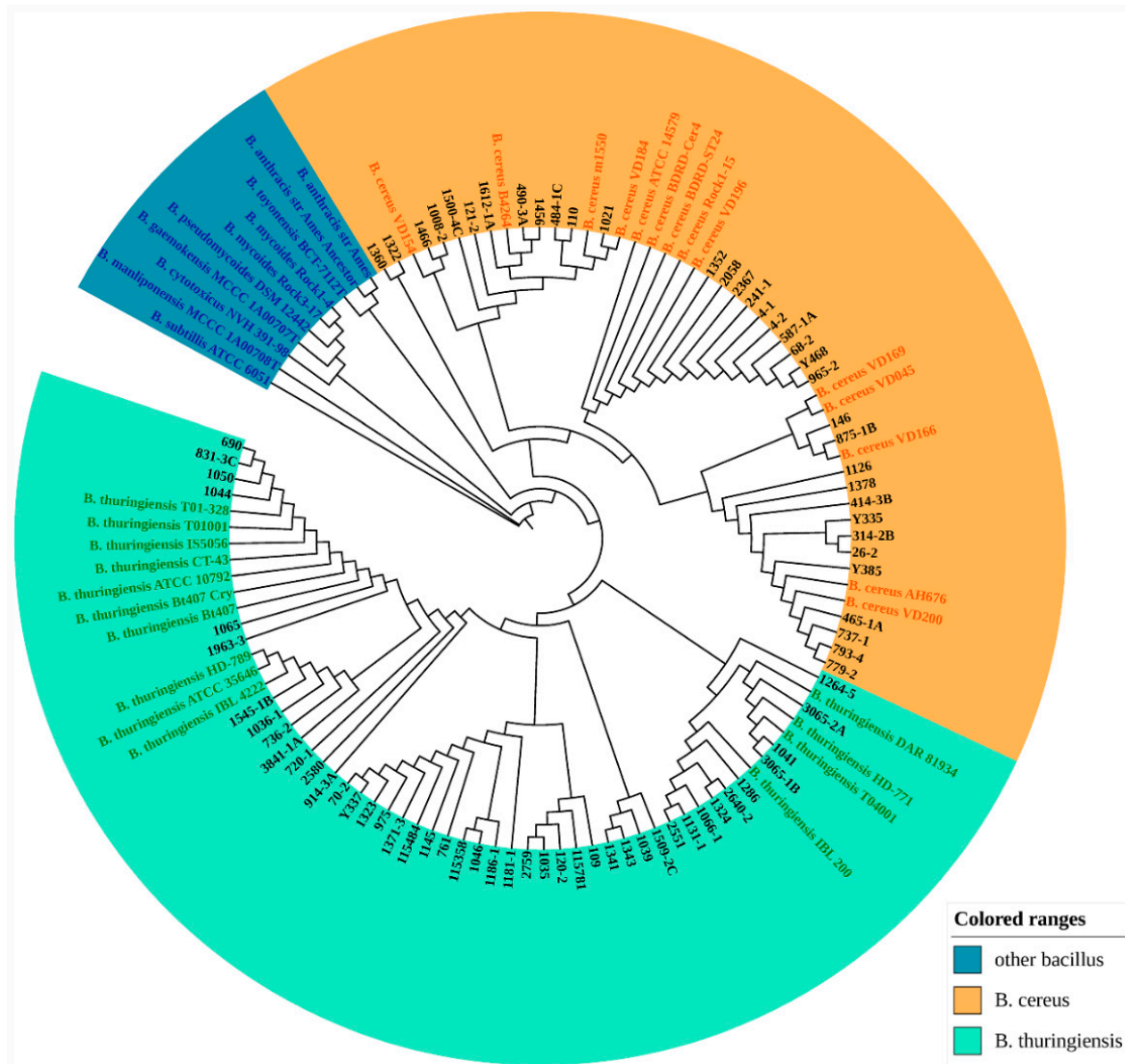

Figure 1. *pycA* gene-based phylogenetic tree of *Bacillus cereus*, *Bacillus thuringiensis*, and other *Bacillus cereus* group species. Branch quality was evaluated using 1000 bootstrap replicates. The reference strains are indicated in color, and the isolated strains are indicated in black.

3.3. Discovery and Identification of Biomarkers in MALDI-TOF MS

Comparative analysis of mass patterns showed that *B. cereus* and *B. thuringiensis* each possessed four species-specific peaks (Figure 2), which were highly conserved and could be considered as species-level biomarkers (Table 2). The peaks at m/z 6427 and 9214 were reproducibly found in 100% (36 of 36) *B. cereus* strains in the present study, and the peaks at m/z 3211 and 9188 were observed in 97.2% of the MS acquisitions. Meanwhile, the peaks at m/z 3218, 6441, 9160, and 9229 were detected in 100% (45 of 45) *B. thuringiensis* strains used in the present study.
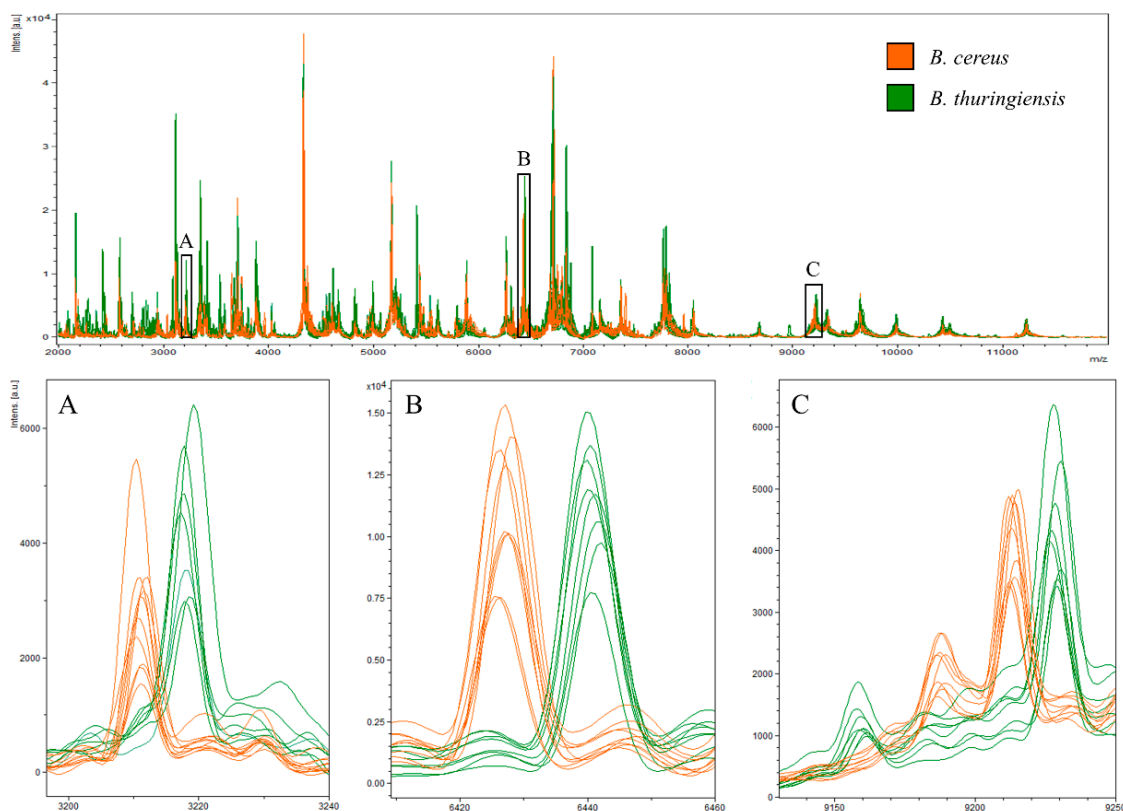
Figure 2. MALDI-TOF mass spectra and species-specific peaks of *Bacillus cereus* (orange) and *Bacillus thuringiensis* (green). The *y*-axis indicates peak intensities, and the *x*-axis indicates the m/z values; (**A**–**C**) represent enlarged views of species-specific peaks.

Table 2. Frequencies and assignments of species-specific peaks for *Bacillus cereus* and *Bacillus thuringiensis*.

Experimental m/z	Presence of Peak (%)		Protein Name	Post-Translational Modification	Theoretical m/z	Amino Acid Substitution
	B. cereus	*B. thuringiensis*				
3211	97.22 (35/36)	0.00 (0/45)	50S ribosomal protein L30 *	methionine removed	3213	V → L
3218	0.00 (0/36)	100.00 (45/45)			3220	
6427	100.00 (36/36)	0.00 (0/45)	50S ribosomal protein L30		6426	
6441	0.00 (0/36)	100.00 (45/45)			6440	
9160	0.00 (0/36)	100.00 (45/45)	50S ribosomal protein L31 type B	-	9158	N → S, L → I
9188	97.22 (35/36)	0.00 (0/45)			9185	
9214	100.00 (36/36)	0.00 (0/45)	30S ribosomal protein S20	methionine removed	9212	A → S
9229	0.00 (0/36)	100.00 (45/45)			9228	

* doubly charged ions.

By comparing the experimental m/z with genome data, the ion peaks at m/z 3211 and 3218 m/z were putatively identified as double-charged ions of the 50S ribosomal protein L30 at m/z 6427 and 6441. The peaks at m/z 9160 and 9188 were assigned to 50S ribosomal protein L31 type B, and the peaks at m/z 9214 and 9229 were characterized as 30S ribosomal protein S20. While a fraction of *B. cereus* lacked the peak at m/z 9188 representing 50S ribosomal protein L31 type B in the MALDI-TOF MS measurement (Table 2), the same result was consistently observed in the genome analysis (Table 1). Furthermore, the sequences of corresponding genes revealed non-synonymous mutations resulting amino acid substitutions in corresponding proteins, producing peak shifts from *B. cereus* to *B. thuringiensis* (Table 2).

3.4. Assessment of Practical Application of MALDI-TOF MS

To characterize biomarker specificity, the performance of MALDI-TOF MS targeting four biomarkers was evaluated with 13 common foodborne pathogens. Most of the other foodborne pathogens we tested did not possess the biomarkers owned by *B. cereus* or *B. thuringiensis* (Table 3 and Figure S1). The four biomarkers avoided overlapping with other common foodborne pathogenic strains and therefore achieved high specificity in routine application.

Table 3. *Bacillus* strains and other common foodborne pathogens tested in the present study and MALDI-TOF MS results of specificity tests.

Bacterial Species	Strain	m/z of Biomarkers							
		3211	6427	9188	9214	3218	6441	9160	9229
Bacillus cereus	ATCC 14579	+	+	+	+	−	−	−	−
Bacillus thuringiensis	ATCC 10792	−	−	−	−	+	+	+	+
Bacillus megaterium	ATCC 14581	−	−	−	−	−	−	−	−
Escherichia coli	ATCC 25922	−	−	−	−	−	−	−	+
Escherichia coli	ATCC 8739	−	−	−	−	−	−	−	+
Salmonella Enteritidis	CMCC 50335	−	−	−	−	−	−	−	−
Salmonella Typhimurium	ATCC 14028	−	−	−	−	−	−	−	−
Vibrio parahaemolyticus	ATCC 33847	+	+	−	−	−	−	−	−
Listeria monocytogenes	ATCC 19115	−	−	−	−	−	−	−	−
Staphylococcus aureus	ATCC 25923	+	+	−	−	−	−	−	−
Cronobacter sakazakii	ATCC 29544	−	−	−	−	−	−	−	−
Pseudomonas aeruginosa	ATCC 15442	−	−	−	−	−	−	−	−
Campylobacter jejuni	ATCC 33291	−	−	−	−	+	−	+	−
Yersinia enterocolitica	CMCC 52204	−	−	−	−	−	−	−	−
Klebsiella pneumoniae	ATCC 700603	−	−	−	−	−	−	−	−

+/− indicate the positive and negative results, respectively.

4. Discussion

Genome-based taxonomy has become increasingly important in bacterial taxonomy and is now recognized as a more standardized taxonomic framework based on robust phylogenetic analysis than traditional classification methods [21]. However, genome-based taxonomy is time-consuming, expensive, and labor-intensive; thus, it cannot be used for fast and high-throughput microbial identification. As MALDI-TOF MS is a fast, cost-efficient, and robust approach for microbial identification, in the present study, we screened the correct *B. cereus* and *B. thuringiensis* using the GTDB and rMLST databases and transformed the robust genome variations (especially variations in ribosomal protein genes) into visible peak differences in MALDI-TOF MS, providing an easy and rapid tool for differentiating between *B. cereus* and *B. thuringiensis*. This method greatly shortened the detection time (Figure 3) and enabled high accuracy.
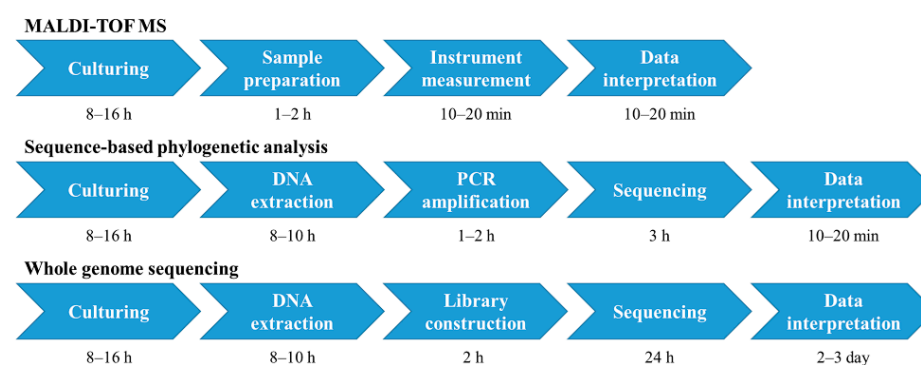

Figure 3. Procedures and time requirements in different methods. The time cost was calculated based on the preparation of 96 samples.

Several studies have been conducted on the applicability of MALDI-TOF MS for discriminating strains within *B. cereus* group. MALDI-TOF MS combined with statistical method and classifying model showing great clustering for *B. anthracis*, *B. cereus*, *B. mycoides*, *B. wiedmannii*, *B. thuringiensis*, *B. toyonensis*, and *B. weihenstephanensis* [33]. Small molecules secreted by *B. cereus* group also showed the discriminatory power, Ha pointed out that m/z at 714.2 and 906.5 can be potential biomarkers to differentiate *B. cereus* and *B. thuringiensis* [32]. Detection of specific mass peaks in the low-mass range may requires better performing instrument, which is not available in every laboratory. In the present study, we utilized the wealth of the genome data to identify ribosomal proteins that are useful for taxonomic identification, resulting in clear differentiation of *B. cereus* and *B. thuringiensis*. The species-specific peaks detected in our study were also observed in other studies, such as m/z at 6427 ± 3 and 9214 ± 3 for *B. cereus*, m/z at 6441 ± 3 and 9229 ± 3 for *B. thuringiensis* [50], showing the inter-laboratory reproducibility of this method. As one biomarker was insufficient for high-resolution taxonomy, we combined three ribosomal proteins (four specific peaks) to reliably differentiate *B. cereus* and *B. thuringiensis* with high specificity.

In this study, the genomic analysis focused on ribosomal proteins, as the corresponding genes were the basis of molecular taxonomy [51]. Conserved ribosomal proteins were highly abundant in the cytoplasm and readily detected by MALDI-TOF MS, which showed its discriminatory power by detecting slight variations in the mass of ribosomal proteins due to amino acid substitutions. The robust substitutions in the amino acid sequences were derived from non-synonymous mutations in the corresponding genes [52]. The steady accumulation of non-synonymous mutations offered genetic diversity to resolve phylogenetic relationships between isolates [53]. Recent taxonomic studies have revealed that a MALDI-TOF MS-based typing tool showed high concordance with molecular typing methods such as MLST [36,54] and core-genome analysis [34]. Tamura and co-workers developed the *S10*-GERMS (*S10-spc-alpha* operon gene encoded ribosomal protein mass spectrum) method and confirmed ribosomal subunit proteins (S10, S16, S20 and L30) could discriminate psychrotolerant species of the *B. cereus* group [50]. Fiedoruk analyzed the peak masses of *B. cereus* group and observed that ribosomal proteins L31 had the highest discriminative value to differentiate emetic *B. cereus* [55].

The combination of genomics and MALDI-TOF MS measurement not only transferred reliable genome-based taxonomy to a faster MALDI-TOF MS platform but also supported the interpretation of different mass patterns with objective genome data. As ribosomal proteins provide phylogenetic information and are accessible to many bacteria [56], this approach provides a framework to identify other closely related species. In the future, MALDI-TOF MS has the potential to differentiate other closely related species and phylogenetically-related strains with different levels of virulence [57], antibiotic resistance [58,59], and host origin [34].

5. Conclusions

By combining genomics and MALDI-TOF MS measurement, we identified four specific mass peaks for differentiating *B. cereus* and *B. thuringiensis* with high specificity. Genome-guided MALDI-TOF MS is suitable for high-throughput, cost-efficient, and reliable differentiation of *B. cereus* and *B. thuringiensis*; could serve as a first-line identification tool, and can be extended to other closely related species.

References

1. EFSA Panel on Biological Hazards (BIOHAZ). Risks for public health related to the presence of *Bacillus cereus* and other *Bacillus* spp. Including *Bacillus thuringiensis* in foodstuffs. *EFSA J.* **2016**, *14*, 4524. [CrossRef]
2. Stenfors Arnesen, L.P.; Fagerlund, A.; Granum, P.E. From soil to gut: *Bacillus cereus* and its food poisoning toxins. *FEMS Microbiol. Rev.* **2008**, *32*, 579–606. [CrossRef] [PubMed]
3. Rouzeau-Szynalski, K.; Stollewerk, K.; Messelhäusser, U.; Ehling-Schulz, M. Why be serious about emetic *Bacillus cereus*: Cereulide production and industrial challenges. *Food Microbiol.* **2020**, *85*, 103279. [CrossRef] [PubMed]
4. Scallan, E.; Hoekstra, R.M.; Angulo, F.J.; Tauxe, R.V.; Widdowson, M.A.; Roy, S.L.; Jones, J.L.; Griffin, P.M. Foodborne illness acquired in the United States—Major pathogens. *Emerg. Infect. Dis.* **2011**, *17*, 7–15. [CrossRef]
5. Paudyal, N.; Pan, H.; Liao, X.; Zhang, X.; Li, X.; Fang, W.; Yue, M. A meta-analysis of major foodborne pathogens in Chinese food commodities between 2006 and 2016. *Foodborne Pathog. Dis.* **2018**, *15*, 187–197. [CrossRef]
6. de Maagd, R.A.; Bravo, A.; Berry, C.; Crickmore, N.; Schnepf, H.E. Structure, diversity, and evolution of protein toxins from spore-forming entomopathogenic bacteria. *Annu. Rev. Genet.* **2003**, *37*, 409–433. [CrossRef] [PubMed]
7. Xu, C.; Wang, B.C.; Yu, Z.; Sun, M. Structural insights into *Bacillus thuringiensis* Cry, Cyt and parasporin toxins. *Toxins* **2014**, *6*, 2732–2770. [CrossRef]
8. Armada, E.; Probanza, A.; Roldán, A.; Azcón, R. Native plant growth promoting bacteria *Bacillus thuringiensis* and mixed or individual mycorrhizal species improved drought tolerance and oxidative metabolism in *Lavandula dentata* plants. *J. Plant Physiol.* **2016**, *192*, 1–12. [CrossRef] [PubMed]
9. Nayak, P.S.; Arakha, M.; Kumar, A.; Asthana, S.; Mallick, B.C.; Jha, S. An approach towards continuous production of silver nanoparticles using *Bacillus thuringiensis*. *RSC Adv.* **2016**, *6*, 8232–8242. [CrossRef]
10. Chen, S.; Deng, Y.; Chang, C.; Lee, J.; Cheng, Y.; Cui, Z.; Zhou, J.; He, F.; Hu, M.; Zhang, L.H. Pathway and kinetics of cyhalothrin biodegradation by *Bacillus thuringiensis* strain ZS-19. *Sci. Rep.* **2015**, *5*, 8784. [CrossRef] [PubMed]
11. Baek, I.; Lee, K.; Goodfellow, M.; Chun, J. Comparative genomic and phylogenomic analyses clarify relationships within and between *Bacillus cereus* and *Bacillus thuringiensis*: Proposal for the recognition of two *Bacillus thuringiensis* genomovars. *Front. Microbiol.* **2019**, *10*, 1978. [CrossRef] [PubMed]
12. Castillo-Esparza, J.F.; Hernández-González, I.; Ibarra, J.E. Search for Cry proteins expressed by *Bacillus* spp. genomes, using hidden Markov model profiles. *3 Biotech* **2019**, *9*, 13. [CrossRef] [PubMed]
13. Méric, G.; Mageiros, L.; Pascoe, B.; Woodcock, D.J.; Mourkas, E.; Lamble, S.; Bowden, R.; Jolley, K.A.; Raymond, B.; Sheppard, S.K. Lineage-specific plasmid acquisition and the evolution of specialized pathogens in *Bacillus thuringiensis* and the *Bacillus cereus* group. *Mol. Ecol.* **2018**, *27*, 1524–1540. [CrossRef] [PubMed]
14. Sacchi, C.T.; Whitney, A.M.; Mayer, L.W.; Morey, R.; Steigerwalt, A.; Boras, A.; Weyant, R.S.; Popovic, T. Sequencing of 16S rRNA gene: A rapid tool for identification of *Bacillus anthracis*. *Emerg. Infect. Dis.* **2002**, *8*, 1117–1123. [CrossRef] [PubMed]
15. Zhong, W.; Shou, Y.; Yoshida, T.M.; Marrone, B.L. Differentiation of *Bacillus anthracis*, *B. cereus*, and *B. thuringiensis* by using pulsed-field gel electrophoresis. *Appl. Environ. Microbiol.* **2007**, *73*, 3446–3449. [CrossRef]
16. Chelliah, R.; Wei, S.; Park, B.J.; Rubab, M.; Banan-Mwine Dalirii, E.; Barathikannan, K.; Jin, Y.G.; Oh, D.H. Whole genome sequence of *Bacillus thuringiensis* ATCC 10792 and improved discrimination of *Bacillus thuringiensis* from *Bacillus cereus* group based on novel biomarkers. *Microb. Pathog.* **2019**, *129*, 284–297. [CrossRef] [PubMed]

17. Park, B.J.; Chelliah, R.; Wei, S.; Park, J.H.; Forghani, F.; Park, Y.S.; Cho, M.S.; Park, D.S.; Oh, D.H. Unique biomarkers as a potential predictive tool for differentiation of *Bacillus cereus* group based on real-time PCR. *Microb. Pathog.* **2018**, *115*, 131–137. [CrossRef]

18. Wei, S.; Chelliah, R.; Park, B.J.; Kim, S.H.; Forghani, F.; Cho, M.S.; Park, D.S.; Jin, Y.G.; Oh, D.H. Differentiation of *Bacillus thuringiensis* from *Bacillus cereus* group using a unique marker based on real-time PCR. *Front. Microbiol.* **2019**, *10*, 883. [CrossRef]

19. Chun, J.; Oren, A.; Ventosa, A.; Christensen, H.; Arahal, D.R.; da Costa, M.S.; Rooney, A.P.; Yi, H.; Xu, X.W.; De Meyer, S.; et al. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 461–466. [CrossRef] [PubMed]

20. Chun, J.; Rainey, F.A. Integrating genomics into the taxonomy and systematics of the *Bacteria* and *Archaea*. *Int. J. Syst. Evol. Microbiol.* **2014**, *64*, 316–324. [CrossRef] [PubMed]

21. Parks, D.H.; Chuvochina, M.; Waite, D.W.; Rinke, C.; Skarshewski, A.; Chaumeil, P.A.; Hugenholtz, P. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **2018**, *36*, 996–1004. [CrossRef] [PubMed]

22. Auch, A.F.; von Jan, M.; Klenk, H.P.; Göker, M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand. Genom. Sci.* **2010**, *2*, 117–134. [CrossRef] [PubMed]

23. Liu, Y.; Lai, Q.; Göker, M.; Meier-Kolthoff, J.P.; Wang, M.; Sun, Y.; Wang, L.; Shao, Z. Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Sci. Rep.* **2015**, *5*, 14082. [CrossRef] [PubMed]

24. Wattal, C.; Oberoi, J.K.; Goel, N.; Raveendran, R.; Khanna, S. Matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS) for rapid identification of micro-organisms in the routine clinical microbiology laboratory. *Eur. J. Clin. Microbiol. Infect. Dis.* **2017**, *36*, 807–812. [CrossRef] [PubMed]

25. Jadhav, S.; Gulati, V.; Fox, E.M.; Karpe, A.; Beale, D.J.; Sevior, D.; Bhave, M.; Palombo, E.A. Rapid identification and source-tracking of *Listeria monocytogenes* using MALDI-TOF mass spectrometry. *Int. J. Food Microbiol.* **2015**, *202*, 1–9. [CrossRef] [PubMed]

26. Manukumar, H.M.; Umesha, S. MALDI-TOF-MS based identification and molecular characterization of food associated methicillin-resistant *Staphylococcus aureus*. *Sci. Rep.* **2017**, *7*, 11414. [CrossRef] [PubMed]

27. Sala-Comorera, L.; Vilaró, C.; Galofré, B.; Blanch, A.R.; García-Aljaro, C. Use of matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry for bacterial monitoring in routine analysis at a drinking water treatment plant. *Int. J. Hyg. Environ. Health* **2016**, *219*, 577–584. [CrossRef] [PubMed]

28. Mangmee, S.; Reamtong, O.; Kalambaheti, T.; Roytrakul, S.; Sonthayanon, P. MALDI-TOF mass spectrometry typing for predominant serovars of non-typhoidal *Salmonella* in a Thai broiler industry. *Food Control* **2020**, *113*, 107188. [CrossRef]

29. Ojima-Kato, T.; Yamamoto, N.; Suzuki, M.; Fukunaga, T.; Tamura, H. Discrimination of *Escherichia coli* O157, O26 and O111 from other serovars by MALDI-TOF MS based on the S10-GERMS method. *PLoS ONE* **2014**, *9*, e113458. [CrossRef] [PubMed]

30. Ojima-Kato, T.; Yamamoto, N.; Takahashi, H.; Tamura, H. Matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) can precisely discriminate the lineages of *Listeria monocytogenes* and species of *Listeria*. *PLoS ONE* **2016**, *11*, e0159730. [CrossRef]

31. Camoez, M.; Sierra, J.M.; Dominguez, M.A.; Ferrer-Navarro, M.; Vila, J.; Roca, I. Automated categorization of methicillin-resistant *Staphylococcus aureus* clinical isolates into different clonal complexes by MALDI-TOF mass spectrometry. *Clin. Microbiol. Infect.* **2016**, *22*, 161.E1–161.E7. [CrossRef] [PubMed]

32. Ha, M.; Jo, H.; Choi, E.; Kim, Y.; Kim, J.; Cho, H. Reliable identification of *Bacillus cereus* group species using low mass biomarkers by MALDI-TOF MS. *J. Microbiol. Biotechnol.* **2019**, *29*, 887–896. [CrossRef] [PubMed]

33. Manzulli, V.; Rondinone, V.; Buchicchio, A.; Serrecchia, L.; Cipolletta, D.; Fasanella, A.; Parisi, A.; Difato, L.; Iatarola, M.; Aceti, A.; et al. Discrimination of *Bacillus cereus* group members by MALDI-TOF mass spectrometry. *Microorganisms* **2021**, *9*, 1202. [CrossRef] [PubMed]

34. Rothen, J.; Pothier, J.F.; Foucault, F.; Blom, J.; Nanayakkara, D.; Li, C.; Ip, M.; Tanner, M.; Vogel, G.; Pflüger, V.; et al. Subspecies typing of *Streptococcus agalactiae* based on ribosomal subunit protein mass variation by MALDI-TOF MS. *Front. Microbiol.* **2019**, *10*, 471. [CrossRef] [PubMed]

35. Corver, J.; Sen, J.; Hornung, B.V.H.; Mertens, B.J.; Berssenbrugge, E.K.L.; Harmanus, C.; Sanders, I.M.J.G.; Kumar, N.; Lawley, T.D.; Kuijper, E.J.; et al. Identification and validation of two peptide markers for the recognition of *Clostridioides difficile* MLST-1 and MLST-11 by MALDI-MS. *Clin. Microbiol. Infect.* **2019**, *25*, 904.E1–904.E7. [CrossRef] [PubMed]

36. Emele, M.F.; Možina, S.S.; Lugert, R.; Bohne, W.; Masanta, W.O.; Riedel, T.; Groß, U.; Bader, O.; Zautner, A.E. Proteotyping as alternate typing method to differentiate *Campylobacter coli* clades. *Sci. Rep.* **2019**, *9*, 4244. [CrossRef] [PubMed]

37. Jolley, K.A.; Bliss, C.M.; Bennett, J.S.; Bratcher, H.B.; Brehony, C.; Colles, F.M.; Wimalarathna, H.; Harrison, O.B.; Sheppard, S.K.; Cody, A.J.; et al. Ribosomal multilocus sequence typing: Universal characterization of bacteria from domain to strain. *Microbiology* **2012**, *158*, 1005–1015. [CrossRef]

38. Gao, T.; Ding, Y.; Wu, Q.; Wang, J.; Zhang, J.; Yu, S.; Yu, P.; Liu, C.; Kong, L.; Feng, Z.; et al. Prevalence, virulence genes, antimicrobial susceptibility, and genetic diversity of *Bacillus cereus* isolated from pasteurized milk in China. *Front. Microbiol.* **2018**, *9*, 533. [CrossRef]

39. Guo, H.; Yu, P.; Yu, S.; Wang, J.; Zhang, J.; Zhang, Y.; Liao, X.; Wu, S.; Ye, Q.; Yang, X.; et al. Incidence, toxin gene profiling, antimicrobial susceptibility, and genetic diversity of *Bacillus cereus* isolated from quick-frozen food in China. *LWT-Food Sci. Technol.* **2021**, *140*, 110824. [CrossRef]

40. Liu, C.; Yu, P.; Yu, S.; Wang, J.; Guo, H.; Zhang, Y.; Zhang, J.; Liao, X.; Li, C.; Wu, S.; et al. Assessment and molecular characterization of *Bacillus cereus* isolated from edible fungi in China. *BMC Microbiol.* **2020**, *20*, 310. [CrossRef]

41. Yu, S.; Yu, P.; Wang, J.; Li, C.; Guo, H.; Liu, C.; Kong, L.; Yu, L.; Wu, S.; Lei, T.; et al. A study on prevalence and characterization of *Bacillus cereus* in ready-to-eat foods in China. *Front. Microbiol.* **2020**, *10*, 3043. [CrossRef] [PubMed]

42. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [CrossRef] [PubMed]

43. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T.; Fookes, M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **2015**, *31*, 3691–3693. [CrossRef] [PubMed]

44. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef] [PubMed]

45. Fasolato, L.; Cardazzo, B.; Carraro, L.; Fontana, F.; Novelli, E.; Balzan, S. Edible processed insects from e-commerce: Food safety with a focus on the *Bacillus cereus* group. *Food Microbiol.* **2018**, *76*, 296–303. [CrossRef] [PubMed]

46. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. Mega X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef] [PubMed]

47. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [CrossRef] [PubMed]

48. Gibb, S.; Strimmer, K. MALDIquant: A versatile R package for the analysis of mass spectrometry data. *Bioinformatics* **2012**, *28*, 2270–2271. [CrossRef]

49. López-Fernández, H.; Santos, H.M.; Capelo, J.L.; Fdez-Riverola, F.; Glez-Peña, D.; Reboiro-Jato, M. Mass-Up: An all-in-one open software application for MALDI-TOF mass spectrometry knowledge discovery. *BMC Bioinform.* **2015**, *16*, 318. [CrossRef]

50. Takahashi, N.; Nagai, S.; Fujita, A.; Ido, Y.; Kato, K.; Saito, A.; Moriya, Y.; Tomimatsu, Y.; Kaneta, N.; Tsujimoto, Y.; et al. Discrimination of psychrotolerant *Bacillus cereus* group based on MALDI-TOF MS analysis of ribosomal subunit proteins. *Food Microbiol.* **2020**, *91*, 103542. [CrossRef]

51. Welker, M.; Van Belkum, A.; Girard, V.; Charrier, J.P.; Pincus, D. An update on the routine application of MALDI-TOF MS in clinical microbiology. *Expert Rev. Proteom.* **2019**, *16*, 695–710. [CrossRef] [PubMed]

52. Grenga, L.; Pible, O.; Armengaud, J. Pathogen proteotyping: A rapidly developing application of mass spectrometry to address clinical concerns. *Clin. Mass Spectrom.* **2019**, *14*, 9–17. [CrossRef]

53. Lieberman, T.D.; Michel, J.B.; Aingaran, M.; Potter-Bynoe, G.; Roux, D.; Davis, M.R., Jr.; Skurnik, D.; Leiby, N.; LiPuma, J.J.; Goldberg, J.B.; et al. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nat. Genet.* **2011**, *43*, 1275–1280. [CrossRef] [PubMed]

54. Meng, X.; Yang, J.; Duan, J.; Liu, S.; Huang, X.; Wen, X.; Huang, X.; Fu, C.; Li, J.; Dou, Q.; et al. Assessing molecular epidemiology of carbapenem-resistant *Klebsiella pneumoniae* (CR-KP) with MLST and MALDI-TOF in central China. *Sci. Rep.* **2019**, *9*, 2271. [CrossRef]

55. Fiedoruk, K.; Daniluk, T.; Fiodor, A.; Drewicka, E.; Buczynska, K.; Leszczynska, K.; Bideshi, D.K.; Swiecicka, I. MALDI-TOF MS portrait of emetic and non-emetic *Bacillus cereus* group members. *Electrophoresis* **2016**, *37*, 2235–2247. [CrossRef] [PubMed]

56. Moore, K.R.; Magnabosco, C.; Momper, L.; Gold, D.A.; Bosak, T.; Fournier, G.P. An expanded ribosomal phylogeny of Cyanobacteria supports a deep placement of plastids. *Front. Microbiol.* **2019**, *10*, 1612. [CrossRef]

57. Emele, M.F.; Joppe, F.M.; Riedel, T.; Overmann, J.; Rupnik, M.; Cooper, P.; Kusumawati, R.L.; Berger, F.K.; Laukien, F.; Zimmermann, O.; et al. Proteotyping of *Clostridioides difficile* as alternate typing method to ribotyping is able to distinguish the ribotypes RT027 and RT176 from other ribotypes. *Front. Microbiol.* **2019**, *10*, 2087. [CrossRef]

58. Flores-Treviño, S.; Garza-González, E.; Mendoza-Olazarán, S.; Morfín-Otero, R.; Camacho-Ortiz, A.; Rodríguez-Noriega, E.; Martínez-Meléndez, A.; Bocanegra-Ibarias, P. Screening of biomarkers of drug resistance or virulence in ESCAPE pathogens by MALDI-TOF mass spectrometry. *Sci. Rep.* **2019**, *9*, 18945. [CrossRef] [PubMed]

59. Wang, H.Y.; Chung, C.R.; Wang, Z.; Li, S.; Chu, B.Y.; Horng, J.T.; Lu, J.J.; Lee, T.Y. 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.* **2020**, *22*, bbaa138. [CrossRef] [PubMed]