Journal of Microbiological Methods 119 (2015) 233–238

Contents lists available at ScienceDirect

Journal of Microbiological Methods

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

Assessing the performance of novel software Strain Solution on automated discrimination of *Escherichia coli* serotypes and their mixtures using matrix-assisted laser desorption ionization-time of flight mass spectrometry

Teruyo Ojima-Kato [a,*], Naomi Yamamoto [b], Yoshio Iijima [c], Hiroto Tamura [a,b,**]

[a] Knowledge Hub Aichi, Aichi Science and Technology Foundation, Yakusa, Toyota, Aichi 470-0356, Japan
[b] School of Agriculture, Meijo University, Shiogamaguchi, Tenpaku-ku, Nagoya, Aichi 468-0073, Japan
[c] Kobe Institute of Health, Minatojimanakamachi, Chuo-ku, Kobe, Hyogo 650-0046, Japan

ARTICLE INFO

Article history:
Received 28 July 2015
Received in revised form 5 November 2015
Accepted 7 November 2015
Available online 10 November 2015

Keywords:
MALDI-TOF MS
Strain Solution
Escherichia coli
S10-GERMS

ABSTRACT

O157, O26, and O111 are the most important O serogroups of enterohemorrhagic *Escherichia coli* worldwide. Recently we reported a strategy for discriminating these serotypes from the others using matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) based on the *S10-spc-alpha* operon gene-encoded ribosomal protein mass spectrum (*S10*-GERMS) method. To realize the fully automated identification of microorganisms at species- or serotype-level with the concept of *S10*-GERMS method, novel software named Strain Solution for MALDI-TOF MS was developed. In this study, the Strain Solution was evaluated with a total of 45 *E. coli* isolates including O26, O91, O103, O111, O115, O121, O128, O145, O157, O159, and untyped serotypes. The Strain Solution could accurately discriminate 92% (11/12) of O157 strains, 100% (13/13) of O26 and O111 strains from the others with three biomarkers in an automated manner. In addition, this software could identify 2 different *E. coli* strains (K-12 as a non-O157 representative and O157) in mixed samples. The results suggest that Strain Solution will be useful for species- or serotype-level classification of microorganisms in the fields of food safety and diagnostics.

1. Introduction

Matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS) has been increasingly used in recent years for the rapid and reliable identification of microorganisms. MALDI-TOF MS enables identification of an unknown bacterial colony within a few minutes without the use of special techniques or costly reagents (Claydon et al., 1996; Cherkaoui et al., 2010). The conventional determinations are based on fingerprinting analysis to compare MS patterns of unknown samples with reference spectral collections in accumulated databases (Jarman et al., 2000; Kallow et al., 2000). MALDI-TOF MS analysis has great potential as an advanced identification method for microorganisms in routine laboratory settings because of the significantly reduced times compared with the molecular biological method. Data from numerous studies using MALDI-TOF MS have proven the fingerprinting approach such as SARAMIS™ (BioMérieux) or Biotyper (Bruker Daltonics) to be powerful for bacterial identification at the genus level. However, serotype- or species-level identification is not always successful because slight mass-to-charge ratio [*m*/*z*] differences in individual MS peaks are not taken into account in fingerprinting method.

In recent years MALDI-TOF MS-based species- or serotype-level typing methods have been broadly studied (Benagli et al., 2012; Christner et al., 2014. Josten et al., 2013; Khot and Fisher, 2013). The *S10-spc-alpha* operon gene-encoded ribosomal protein mass spectrum (*S10*-GERMS) method was developed to enable database searching with MALDI-TOF MS results and reliable species- or subspecies-level identification of microorganisms (Hotta et al., 2010). When using MALDI-TOF MS data to identify microorganisms, most of high-intensity MS peaks detected are derived from ribosomal proteins encoded in the *S10-spc-alpha* operon, where over half of the ribosomal subunit proteins are encoded. The *S10*-GERMS method provides calculated theoretical [*m*/*z*] values based on DNA sequences of genes in the

Abbreviations: MALDI-TOF MS, matrix-assisted laser desorption ionization-time of flight mass spectrometry; *S10*-GERMS, *S10-spc-alpha* operon gene-encoded ribosomal protein mass spectrum; EHEC, enterohemorrhagic *E. coli*; CHCA, α-cyano-4-hydroxycinnamic acid; SA, sinapic acid.

* Corresponding author.
** Correspondence to: H. Tamura, School of Agriculture, Meijo University, Shiogamaguchi, Tenpaku-ku, Nagoya, Aichi 468-0073, Japan.
E-mail addresses: teruyo.ojima@gmail.com (T. Ojima-Kato), hiroto@meijo-u.ac.jp (H. Tamura).

http://dx.doi.org/10.1016/j.mimet.2015.11.005

S10-spc-alpha operon therefore an accurate typing of the *m/z* of biomarker peaks observed in MALDI-TOF MS can be performed; i.e. "Proteotyping". The *S10*-GERMS method has been effectively employed in the proteotyping of *Bacillus subtilis* and *Pseudomonas syringae* serotypes (or pathovars) (Hotta et al., 2010; Tamura et al., 2013) and *Lactobacillus casei* strains (Sato et al., 2012). Recently we found 4 specific biomarker peaks to discriminate *E. coli* serotypes O157, O26, and O111 from the others. The biomarker was backed by the scientific evidences from 106 *E. coli* strains with 12 different serotypes (Ojima-Kato et al., 2014). Using these specific biomarker peaks detected in MALDI-TOF MS, a semi-automated discrimination was demonstrated with more than a hundred *E. coli* strains. This approach involves a combined analysis of the observed MS differences between each serotype's biomarkers (ribosomal proteins S15, L25 and DNA-binding protein H-NS) and the presence or absence of MS peaks corresponding to stress protein HdeB.

To widely disseminate this rapid and reliable species- or serotype-level microbial discrimination method, both the standardization of the *S10*-GERMS method and the development of convenient software are important. Therefore, our group has developed the analyzing software 'Strain Solution' that is applicable for a routine diagnostic use in microbial laboratories. This software is now available from Shimadzu Corporation (Kyoto, Japan) for the powerful bacterial discrimination tool of AXIMA MALDI-TOF MS system. The concept of the Strain Solution, was designed to accommodate 1) registering the MS *m/z* $[M + H]^+$ peaks of biomarkers generated using the *S10*-GERMS method, 2) importing data matrixes of *m/z* peaks and intensities from an isolated microorganism(s) obtained by MALDI-TOF MS analysis, and 3) matching the registered MS peaks with that of analyzed microorganism(s).

In this report, we evaluate the newly developed software Strain Solution with 45 *E. coli* strains stored in the Kobe Institute of Health which including O26, O103, O91, O111, O115, O121, O128, O145, O157, O159 and untyped serotypes isolated from patients in Japan, USA, and Kenya; and food. In addition, we evaluate Strain Solution to characterize the mixtures of closely related bacteria by using their own biomarker peaks specific to the corresponding strains. The results suggest that Strain Solution will be a useful tool for species- or serotype-level discrimination of microorganisms in the fields of food safety and diagnostics.

2. Materials and methods

2.1. Bacterial strains

Forty-five *E. coli* isolated strains stored in the Kobe Institute of Health were used in this study (Table 1). Shiga toxin 1 (Stx1) and Shiga toxin 2 (Stx2) production was confirmed by the latex coagulating method (Cook, 1965) using the VTEC-RPLA Seiken test (Denka Seiken, Tokyo, Japan) and by PCR using specific primers for detecting VTs, as recommended in the enterohemorrhagic *E. coli* (EHEC) diagnostics manual published by the National Institute of Infectious Diseases in Japan (Lin et al., 1993; Cebula et al., 1995). O-antigen was also determined using antisera against pathogenic *E. coli* (purchased from Denka Seiken) and PCR amplification of the *rfb* O157 gene (420-bp amplicon) (Maurer et al., 1999). The virulence factor *eae* gene was detected by amplifying a 151-bp PCR product, as previously described (Kawasaki et al., 2005). The carriage of the *fliCH7* gene in the KB 0341-2 strain was analyzed according to a previous study (Gannon et al., 1997). PCR products were detected with ethidium bromide straining after running on 2% agarose gel.

Table 1
E. coli strains used in this study.

ID	Serotype	Gene			Source
		stx1	*stx2*	*eae*	
KB0137	O157	+	+	N. D.	Patient (Japan)
KB0139	O157	+	+	N. D.	Patient (Japan)
KB0150	O157	−	−	N. D.	Patient (USA)
KB0152	O157	+	+	N. D.	Patient (Japan)
KB0155	O157	−	+	N. D.	Patient (Kenya)
KB0156	O157	+	+	N. D.	Patient (Japan)
KB0340	O157	−	+	N. D.	Beef liver
KB0341-1	O157	−	+	+	Beef omasum
KB0341-2	O157	−	−	−	Beef omasum
KB0514	O157	−	+	+	Patient (Japan)
KB0521	O111	+	−	+	Patient (Japan)
KB0522	O157	+	+	+	Patient (Japan)
KB0548	O159	+	+	N. D.	Patient (Kenya)
KB0549	UT	−	+	N. D.	Patient (Kenya)
KB0617	O26	+	+	N. D.	Patient (unknown)
KB0618	O26	+	+	N. D.	Patient (unknown)
KB0619	O26	+	+	N. D.	Patient (unknown)
KB0620	O26	+	+	N. D.	Patient (unknown)
KB0621	O26	+	+	N. D.	Patient (unknown)
KB0622	O26	+	+	N. D.	Patient (unknown)
KB0623	O26	+	+	N. D.	Patient (unknown)
KB0624	O26	+	+	N. D.	Patient (unknown)
KB0625	O26	+	+	N. D.	Patient (unknown)
KB0626	O26	+	+	N. D.	Patient (unknown)
KB0627	O111	+	−	N. D.	Patient (Japan)
KB0628	O111	+	+	N. D.	Patient (Japan)
KB0732	O121	−	+	+	Patient (Japan)
KB0733	UT	+	+	−	Minced meat
KB0734	O145	+	−	+	Patient (Japan)
KB0735	O91	+	−	−	Patient (Japan)
KB0738	UT	−	+	+	Patient (Japan)
KB0739	O121	−	+	+	Patient (Japan)
KB0740	O91	+	+	−	Patient (Japan)
KB0741	O145	+	−	+	Patient (Japan)
KB0742	O128	+	+	−	Patient (Japan)
KB0743	O91	+	+	−	Patient (Japan)
KB0744	UT	−	+	−	Patient (Japan)
KB0745	O157	+	+	+	Patient (Japan)
KB0746	O115	+	−	−	Patient (Japan)
KB0747	O121	−	+	+	Patient (Japan)
KB0748	O145	−	+	+	Patient (Japan)
KB0749	O103	+	−	+	Patient (Japan)
KB0750	O103	+	−	+	Patient (Japan)
KB0751	O145	+	−	+	Patient (Japan)
KB0752	O145	−	+	+	Patient (Japan)

Serotypes were determined by performing antisera agglutination tests and PCR.
UT, untyped.
N. D., not determined.

2.2. Cultivation of bacteria and MALDI-TOF MS analysis

Bacterial colonies grown at 37 °C for 20 h on tryptone soy agar plates were analyzed with the AXIMA MALDI-TOF MS System for microorganisms (Shimadzu Corporation) and the SARAMIS database. Briefly, single colonies were spread on analytical metal plates with sterile toothpicks for MALDI-TOF MS analysis. Next, 1 μL of saturated α-cyano-4-hydroxycinnamic acid (CHCA) solution in 50% (v/v) acetonitrile containing 1% trifluoroacetic acid was spotted onto the bacterial cells and mixed well by pipetting. An alternative matrix solution (20 mg/mL of sinapic acid [SA] in the same buffer used to dissolve CHCA) was also examined, as it was recommended for analysis in Strain Solution. In cases where MALDI-TOF analysis was conducted using SA as the matrix, the analytical metal plates were pre-coated with 1 μL of saturated SA in acetone for stabilizing calibration of the instrument (Önnerfjord et al., 1999). *E. coli* DH5α cells were used to calibrate the AXIMA MALDI-TOF MS System, according to the manufacturer's recommended protocol.

2.3. Analysis with Strain Solution

MALDI-TOF MS *m/z* peaks and peak intensity data obtained using either CHCA or SA as the matrix were imported as ASCII format into Strain Solution, version 1.0.0 (Shimadzu Corporation). Three biomarkers, namely ribosomal protein S15 (*m/z* 10,138.6 or 10,166.6), ribosomal

Table 2
MS m/z $[M + H]^+$ of biomarker proteins used for the discrimination of O157, O26, and O111 in Strain Solution.

	Group		
Biomarker	A	B	C
S15	10,166.6	10,138.6	10,138.6
L25	10,676.4	10,694.4	10,694.4
H-NS	15,409.4	15,425.4	15,409.4

Group A: Typical O157; Group B: O26 and O111 groups; Group C: other serotypes.

protein L25 (m/z 10,676.4 or 10,694.4), and acid stress protein H-NS (m/z 15,409.4 or 15,425.4), were programed in Strain Solution as standard biomarkers for discriminating serotypes O157, O26, and O111 from the other serotypes. The parental m/z $[M + H]^+$ peaks of biomarker proteins for 3 groups of serotypes are shown in Table 2. Group A represents O157 type serotypes, Group B represents O26 and O111 type serotypes, and Group C represents the other serotypes tested. Automated analysis was performed using 800 ppm MS tolerances in Strain Solution.

2.4. Cluster analysis

Binary tables of 3 biomarkers and identity scores for each strain were generated from the Strain Solution results, and cluster analysis was performed as described previously (Ojima-Kato et al., 2014).

2.5. Analysis of mixed bacterial strains

To evaluate the capacity of Strain Solution software for analyzing mixed bacterial samples, *E. coli* K 12 ATCC47076 strain and O157 GTC03904 strain (belonging to Groups C and A, respectively) were cultivated separately overnight in tryptone soy broth and then mixed with various cell ratios. The results obtained in MALDI-TOF MS were imported into Strain Solution and analyzed as the same method described above. ATCC47076 and GTC03904 were purchased from the American Type Culture Collection (Rockville, MD, USA) and the National BioResource Project (Gifu University, Gifu, Japan), respectively.

3. Results

3.1. Analysis by Strain Solution

The carriage of O antigen genes and of *eae* genes, and production of Shiga toxins of a total of 45 *E. coli* strains (1996–2011) stored at the Kobe Institute of Health were confirmed as shown in Table 1. Untyped O-antigen strains did not react with any 50 type antisera available from Denka Seiken.

First, they were analyzed by MALDI-TOF MS with CHCA as the matrix. In case of software SARAMIS based on fingerprinting approach, all strains were identified as '*E. coli*' with >95% identities to the parental reference database. Subsequently, m/z and peak intensities datasets generated with an SA matrix were characterized by Strain Solution to classify them into group A, B, and C (Fig. 1), using group definitions shown in Table 2. As a result, all O157 strains were classified into Group A which represent the typical O157 biomarker pattern, but one O157 (KB0341-2) was into group C due to the lack of specific mutations (mass shifts) in the S15 and L25 proteins. On the other hand, all O26 and O111 were classified into Group B which represents the typical O26/O111 biomarker pattern. Therefore the overall accuracy of the automated analysis was 92% (11/12) for O157 serotypes and 100% (13/13) for O26 and O111 serotypes, using the strains stored in the Kobe Institute of Health.

The analytical results obtained using CHCA were identical to those obtained using SA with most strains. However a few strains, namely KB0521 (O111), KB0733 (UT), KB0734 (O145), KB0735 (O91), KB0740 (O91), and KB0749 (O103) were incorrectly classified due to marginal

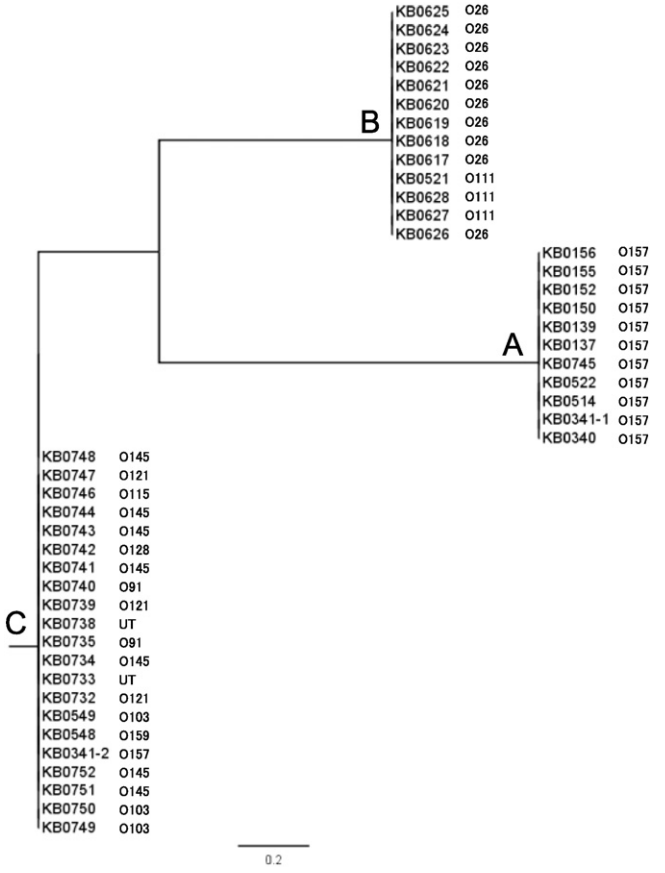

Fig. 1. Cluster analysis of the results obtained using Strain Solution. Designations of A, B, and C indicate group categories according to the 3 biomarker MS patterns listed in Table 2. Group A represents typical O157 type, Group B represents O26 and O111, and Group C represents the other serotypes.

detection of the high molecular weight protein H-NS, which is a defining biomarker specific for O26 and O111 in MALDI-TOF MS analysis (Ojima-Kato et al., 2014). Therefore, the accuracy rates with Strain Solution were 92% (11/12) for O157 and 92% (12/13) for O26 and O111 when using CHCA as the matrix.

3.2. Non-O157-like strains

One O157-type *E. coli* strain (KB 0341-2) classified into group C (other serotypes) by Strain Solution was investigated in greater detail to study correlations between phylogenetic features and the results obtained by Strain Solution. Additional features of strain KB 0341-2 not shown in Table 1 were as follows: carriage of *fliCH7* gene: negative; growth on CT-SMAC: positive; colony color on CT-SMAC agar: red; sorbitol fermentation: positive; and β-glucuronidase activity: positive. Although both agglutination with O157 antisera and carriage of *rfb* O157 gene were positive and it was identified as '*E. coli*' in SARAMIS, the physiological features were different from that of typical O157 strains. This strain did not possess *stx1* or *stx2* genes (Table 1).

3.3. Analysis of mixtures containing 2 different E. coli strains

Since the Strain Solution was successfully employed for serotype identification of *E. coli* with 3 biomarkers, the capabilities of the software in identifying mixtures of 2 different types of *E. coli* strains (K12 as a non-O157 representative and O157) were investigated. First, K12 and O157 strains were classified into group C and A, respectively, with 3 biomarkers using Strain Solution. The analytical results obtained with cell mixtures with various ratios were shown in Table 3. With

Table 3
Detection of biomarkers of mixed samples in Strain Solution.

Ratio	Biomarker hits			Output result from Strain Solution
K12:O157	S15	L25	H-NS	
100:0	C	C	A, C	Other serotypes
90:10	A, C	C	A, C	O157 and other serotypes
80:20	A, C	C	A, C	O157 and other serotypes
70:30	A, C	C	A, C	O157 and other serotypes
60:40	A, C	A, C	A, C	O157 and other serotypes
50:50	A, C	A, C	A, C	O157 and other serotypes
40:60	A, C	A, C	A, C	O157 and other serotypes
30:70	A, C	A, C	A, C	O157 and other serotypes
20:80	A, C	A	A, C	O157 and other serotypes
10:90	A	A	A	O157
0:100	A	A	A	O157

Detections of MS m/z $[M + H]^+$ peaks are indicated as group A (O157 type) and C (other serotypes).

SARAMIS, all samples mixed at various ratios were identified as *E. coli*. In Strain Solution mixed samples containing K12:O157 ratios ranging from 10:90 to 80:20 were successfully recognized that containing both K12 and O157 because the biomarkers corresponding to the registered m/z of S15 (m/z 10,138.6 or 10,166.6) and L25 (m/z 10,676.4 or 10,694.4) were detected with distinct double peaks (Fig. 2).

4. Discussion

MALDI-TOF MS is becoming increasingly used for the rapid, accurate, and low-cost identification of microorganisms. However, the serotype- or strain- level discrimination is difficult by conventional fingerprinting analyzing software. The *S10*-GERMS method was developed with the goal of achieving serotype- or species-level identification by proteotyping that utilizes an accurate biomarker MS database constructed based on combined data from genomics and proteomics approaches (Teramoto et al., 2007; Hotta et al., 2010, 2011, 2012; Tamura et al., 2013; Ojima-Kato et al., 2014). The database can reflect MS differences of even a single amino acid substitution of biomarker peaks which can be caused by specific point mutations in genes in each serotypes or species. The Strain Solution software was developed by our group in collaboration with Shimadzu Corporation to enable automated discrimination of identifications and non-identifications of preset m/z values corresponding to biomarkers that were registered based on the *S10*-GERMS method.

There have been specific requirements for identification of foodborne bacteria *E. coli* at serotype level especially in food microbiological laboratories as well as in clinical and public health laboratories because in particular, *E. coli* serotype O157 is the most common EHEC responsible for a large number of outbreaks in the world. Therefore, we employed *S10*-GERMS method to meet this requirement and successfully discriminated major EHEC serotypes O157, O26, and O111 from the others with a total 106 *E. coli* strains of 12 different serotypes in the previous report (Ojima-Kato et al., 2014). Since the biomarker MS database for *E. coli* was confirmed, in the present study, we utilized this database for the evaluation of the Strain Solution by testing another *E. coli* isolates with various serotypes collected from different countries in 1996–2011. The definitive 3 biomarkers, ribosomal proteins S15, L25, and DNA binding protein H-NS were selected to evaluate the performance of Strain Solution because these biomarker peaks were detected with all *E. coli* analytes ever tested.

From the 45 isolated *E. coli* strains including O26, O103, O91, O111, O115, O121, O128, O145, O157, O159, and untyped serotypes, O157, O26, and O111 were properly discriminated with only 3 biomarkers registered in Strain Solution (Fig. 1, Table 2). Moreover, Strain Solution could accomplish the discrimination of them without any additional parameters for more detailed classification, such as the presence or absence of stress chaperone HdeB (m/z 9066.2) which was specifically absent in serotype O157 (Carter et al., 2012; Ojima-Kato et al., 2014), and the alternative biomarker YdaQ (m/z 8325.6). Analysis with Strain Solution followed by fingerprinting analysis (SARAMIS) in MALDI-TOF MS can be performed continuously, therefore the high-performance solution of bacterial discrimination by MALDI-TOF MS analysis will be achieved as an integrated function in combination with the conventional MALDI-TOF MS fingerprinting software and the Strain Solution. In fact, all samples used in this study were first determined as '*E. coli*' in SARAMIS and then more detailed classification results were obtained from Strain Solution within a few minutes. It suggests that our strategy will be useful for rapid and automated screening of serotypes O157, O26, and O111 directly from a collection of *E. coli*. We believe this sophisticated computerized system is advantageous because conventional serotype determining methods are laborious and errors occur due to misinterpretations of results stemming from the difficulty of discerning fine differences in the colors of bacterial colonies on selective media or in antisera agglutination studies.

While CHCA is mainly used as a matrix for MALDI-TOF MS fingerprinting analysis, SA is recommended for Strain Solution because of its suitability for the detection of high molecular weight proteins. We compared these 2 matrixes for the automated analysis in Strain Solution. The results obtained from either matrix were nearly identical although a few strains other than O157 were not properly classified when using CHCA due to low detectability of H-NS. Therefore, SA is preferable and allows reliable discrimination of *E. coli* with Strain Solution.

The specific m/z shifts of the S15 and L25 in serotype O157 are highly conserved among the genome sequenced *E. coli* O157 strains registered in National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/), implying that our discriminating strategy using Strain Solution with MALDI-TOF MS may be accepted for any O157 strains regardless of the dates or places of isolation. Even the closely related *E. coli* O55, thought to be an ancestor of O157 (Feng et al., 1998; Whittam et al., 1993), can be distinguished because the theoretical

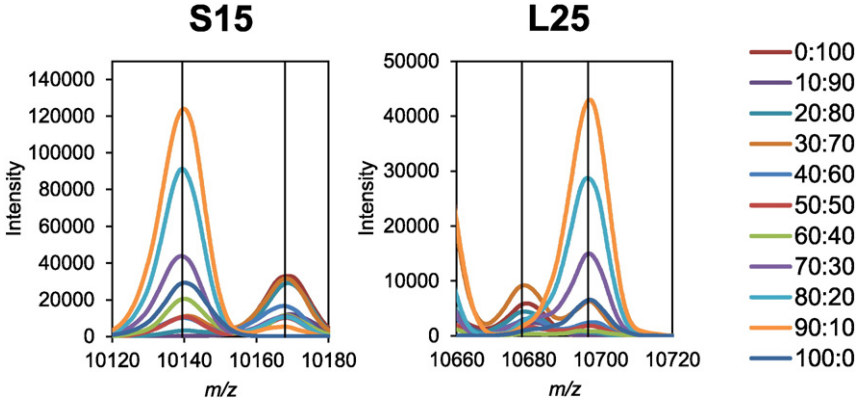

Fig. 2. MS peaks of ribosomal protein S15 and L25 in mixtures containing *E. coli* K12 and O157 strains. The mixed ratio of K12:O157 is referred at the right.

m/*z* values of S15 and L25 are different from those of O157 (data not shown). With regards to O26 and O111, some *E. coli* serotypes such as O118, O69, or O123, whose theoretical *m*/*z* values of H-NS are identical to that of O26 and O111 should be investigated for more detailed classification studies with additional biomarkers.

Only one O157 strain (KB 0341-2) showing abnormal characteristics was misidentified as group C (other serotypes) with Strain Solution due to the lack of specific *m*/*z* shifts for S15 and L25 (Fig. 1). It might be interpreted that Strain Solution could reveal the anomaly of this strain. Among the O157 strains available in NCBI, a small minority (~5%) belongs to group C (other serotypes) for the same reason as mentioned above for KB 0341-2. In addition, a minor group of O157 serotype described as sorbitol-positive, β-glucuronidase-positive, and Shiga toxin-negative has been reported (Díaz et al., 2011). This suggests that the Strain Solution may have a great potential to discriminate the O157 strains with or without *stx1* or *stx2* genes. As yet it is uncertain whether the distinguishing physiological features of this exceptional minor O157 group correlate with the results from Strain Solution, further study will be required.

With the explosion of interest in MALDI-TOF MS bacterial identification, direct MALDI-TOF MS analysis of mixed bacterial samples or crude samples like blood or diarrhea stools of patients is a big matter of concern in clinical laboratories (Kok et al., 2011). However, the available identification methods based on the fingerprinting method do not perform acceptably with mixed bacterial samples (Lagace-Wiens et al., 2012; Fothergill et al., 2013). To overcome such limitations, an advanced algorithm has been reported (Mahé et al., 2014), which may be suitable for discriminating different bacteria at the genus level, but appears to be limited for distinguishing mixtures containing the same species of bacteria. On the other hand, since in the *S10*-GERMS method, the registered biomarkers' *m*/*z* values characteristic to each strain are strictly backed by both the genomics and proteomics approaches, the Strain Solution has great potential to recognize the closely related bacterial mixtures using their own *m*/*z* information specific to the corresponding strains. We evaluated the performance of Strain Solution in identifying mixed bacterial samples containing closely related *E. coli* K12 and O157 strains, which are usually identified as only '*E. coli*' with the conventional fingerprinting method. Experimentally mixed samples were successfully recognized as bacterial mixtures by Strain Solution with 3 discriminative biomarkers (Table 3). Such a capability of Strain Solution will be helpful in monitoring contaminations of similar bacteria for example in food company, public health laboratory or public culture collections that deal with multiple related bacterial samples.

We are going to enlarge accurate MS databases for various microorganisms concerning food industry or human health in order to provide more useful platform for proteotyping using Strain Solution in the field of microbial discrimination.

5. Conclusions

Rapid and automated MALDI-TOF MS analysis with Strain Solution could be effectively applied to discriminate *E. coli* serotypes O157, O26, and O111 which are a major health concern worldwide. The analyzing software Strain Solution based on the accurate MS database constructed by the *S10*-GERMS method will be versatile not only for serotype- or strain-level discrimination but also for recognition of closely related bacterial mixtures in the fields of food safety and diagnostics.

Acknowledgments

This work was financially supported by the Aichi Science and Technology Foundation (Japan) in a part of 'Technological Development Project for Food Safety and Security' under Knowledge Hub Aichi. We would like to thank Mrs. Akiko Ito (Meijo University) and Mrs. Junko Nakayama (Meijo University) for providing technical assistance in sample preparation.

References

Benagli, C., Demarta, A., Caminada, A.P., Ziegler, D., Petrini, O., Tonolla, M., 2012. A rapid MALDI-TOF MS identification database at genospecies level for clinical and environmental *Aeromonas* strains. PLoS One 7, e48441. http://dx.doi.org/10.1371/journal.pone.0048441.

Carter, M.Q., Louie, J.W., Fagerquist, C.K., Sultan, O., Miller, W.G., Mandrell, R.E., 2012. Evolutionary silence of the acid chaperone protein HdeB in enterohemorrhagic *Escherichia coli* O157:H7. Appl. Environ. Microbiol. 78, 1004–1014. http://dx.doi.org/10.1128/AEM.07033-11.

Cebula, T.A., Payne, W.L., Feng, P., 1995. Simultaneous identification of strains of *Escherichia coli* serotype O157:H7 and their Shiga-like toxin type by mismatch amplification mutation assay-multiplex PCR. J. Clin. Microbiol. 33, 248–250.

Cherkaoui, A., Hibbs, J., Emonet, S., Tangomo, M., Girard, M., Francois, P., Schrenzel, J., 2010. Comparison of two matrix-assisted laser desorption ionization-time of flight mass spectrometry methods with conventional phenotypic identification for routine identification of bacteria to the species level. J. Clin. Microbiol. 48, 1169–1175. http://dx.doi.org/10.1128/JCM.01881-09.

Christner, M., Trusch, M., Rohde, H., Kwiatkowski, M., Schlüter, H., Wolters, M., Aepfelbacher, M., Hentschke, M., 2014. Rapid MALDI-TOF mass spectrometry strain typing during a large outbreak of Shiga-toxigenic *Escherichia coli*. PLoS One 9, e101924. http://dx.doi.org/10.1371/journal.pone.0101924.

Claydon, M.A., Davey, S.N., Edwards-Jones, V., Gordon, D.B., 1996. The rapid identification of intact microorganisms using mass spectrometry. Nat. Biotechnol. 14, 1584–1586.

Cook, J.R., 1965. Reversed passive haemagglutination systems for the estimation of tetanus toxins and antitoxins. Immunology 8, 74–80.

Díaz, S., Vidal, D., Herrera-León, S., Sánchez, S., 2011. Sorbitol-fermenting, β-glucuronidase-positive, Shiga toxin-negative *Escherichia coli* O157:H7 in free-ranging red deer in South-Central Spain. Foodborne Pathog. Dis. 8, 1313–1315. http://dx.doi.org/10.1089/fpd.2011.0923.

Feng, P., Lampel, K.A., Karch, H., Whittam, T.S., 1998. Genotypic and phenotypic changes in the emergence of *Escherichia coli* O157:H7. J. Infect. Dis. 177, 1750–1753.

Fothergill, A., Kasinathan, V., Hyman, J., Walsh, J., Drake, T., Wang, Y.F., 2013. Rapid identification of bacteria and yeasts from positive-blood-culture bottles by using a lysis-filtration method and matrix-assisted laser desorption ionization-time of flight mass spectrum analysis with the SARAMIS database. J. Clin. Microbiol. 51, 805–809. http://dx.doi.org/10.1128/JCM.02326-12.

Gannon, P.V., D'Souza, S., Graham, T., King, R.K., Rahn, K., Read, S., 1997. Use of the flagellar H7 gene as a target in multiplex PCR assays and improved specificity in identification of enterohemorrhagic *Escherichia coli* strains. J. Clin. Microbiol. 35, 656–662.

Hotta, Y., Teramoto, K., Sato, H., Yoshikawa, H., Hosoda, A., Tamura, H., 2010. Classification of genus *Pseudomonas* by MALDI-TOF MS based on ribosomal protein coding in *S10-spc-alpha* operon at strain level. J. Proteome Res. 9, 6722–6728. http://dx.doi.org/10.1021/pr100868d.

Hotta, Y., Sato, J., Sato, H., Hosoda, A., Tamura, H., 2011. Classification of the genus *Bacillus* based on MALDI-TOF MS analysis of ribosomal proteins coded in *S10* and *spc* operons. J. Agric. Food Chem. 59, 5222–5230. http://dx.doi.org/10.1021/jf2004095.

Hotta, Y., Sato, H., Hosoda, A., Tamura, H., 2012. MALDI-TOF MS analysis of ribosomal proteins coded in *S10* and *spc* operons rapidly classified *Sphingomonaceae* as alkylphenol polyethoxylate-degrading bacteria from the environment. FEMS Microbiol. Lett. 330, 23–29. http://dx.doi.org/10.1111/j.1574-6968.2012.02525.x.

Jarman, K.H., Cebula, S.T., Saenz, A.J., Petersen, C.E., Valentine, N.B., Kingsley, M.T., Wahl, K.L., 2000. An algorithm for automated bacterial identification using matrix-assisted laser desorption/ionization mass spectrometry. Anal. Chem. 72, 1217–1223.

Josten, M., Reif, M., Szekat, C., Al-Sabti, N., Roemer, T., Sparbier, K., Kostrzewa, M., Rohde, H., Sahl, H.G., Bierbaum, G., 2013. Analysis of the matrix-assisted laser desorption ionization-time of flight mass spectrum of *Staphylococcus aureus* identifies mutations that allow differentiation of the main clonal lineages. J. Clin. Microbiol. 51, 1809–1817. http://dx.doi.org/10.1128/JCM.00518-13.

Kallow W., Dieckmann R., Kleinkauf N., Erhard M., Neuhof, T. (2000) Method of identifying microorganisms using MALDI-TOF MS, European patent, EP1253622.

Kawasaki, S., Horikoshi, N., Okada, Y., Takeshita, K., Sameshima, T., Kawamoto, S., 2005. Multiplex PCR for simultaneous detection of *Salmonella* spp., *Listeria monocytogenes*, and *Escherichia coli* O157:H7 in meat samples. J. Food Prot. 68, 551–556.

Khot, P.D., Fisher, M.A., 2013. Novel approach for differentiating *Shigella* species and *Escherichia coli* by matrix-assisted laser desorption ionization-time of flight mass spectrometry. J. Clin. Microbiol. 51, 3711–3716. http://dx.doi.org/10.1128/JCM.01526-13.

Kok, J., Thomas, L.C., Olma, T., Chen, S.C., Iredell, J.R., 2011. Identification of bacteria in blood culture broths using matrix-assisted laser desorption-ionization Sepsityper and time of flight mass spectrometry. PLoS One 6, e23285. http://dx.doi.org/10.1371/journal.pone.0023285.

Lagace-Wiens, P.R., Adam, H.J., Karlowsky, J.A., Nichol, K.A., Pang, P.F., Guenther, J., Webb, A.A., Miller, C., Alfa, M.J., 2012. Identification of blood culture isolates directly from positive blood cultures by use of matrix-assisted laser desorption ionization time of flight mass spectrometry and a commercial extraction system: analysis of performance, cost, and turnaround time. J. Clin. Microbiol. 50, 3324–3328. http://dx.doi.org/10.1128/JCM.01479-12.

Lin, Z., Kurazono, H., Yamasaki, S., Takeda, Y., 1993. Detection of various variant verotoxin genes in *Escherichia coli* by polymerase chain reaction. Microbiol. Immunol. 37, 543–548.

Mahé, P., Arsac, M., Chatellier, S., Monnin, V., Perrot, N., Mailler, S., Girard, V., Ramjeet, M., Surre, J., Lacroix, B., van Belkum, A., Veyrieras, J.B., 2014. Automatic identification of mixed bacterial species fingerprints in a MALDI-TOF mass-spectrum. Bioinformatics 30, 1280–1286. http://dx.doi.org/10.1093/bioinformatics/btu022.

Maurer, J.J., Schmidt, D., Petrosko, P., Sanchez, S., Bolton, L., Leel, M.D., 1999. Development of primers to O-antigen biosynthesis genes for specific detection of *Escherichia coli* O157 by PCR. Appl. Environ. Microbiol. 65, 2954–2960.

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

Önnerfjord, P., Ekström, S., Bergquist, J., Nilsson, J., Laurell, T., Marko-Verga, G., 1999. Homogeneous sample preparation for automated high throughput analysis with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. 13, 315–322.

Sato, H., Torimura, M., Kitahara, M., Ohkuma, M., Hotta, Y., Tamura, H., 2012. Characterization of the *Lactobacillus casei* group based on the profiling of ribosomal proteins coded in S10-spc-alpha operons as observed by MALDI-TOF MS. Syst. Appl. Microbiol. 35, 447–454. http://dx.doi.org/10.1016/j.syapm.2012.08.008.

Tamura, H., Hotta, Y., Sato, H., 2013. Novel accurate bacterial discrimination by MALDI-time-of-flight MS based on ribosomal proteins coding in *S10-spc-alpha* operon at strain level *S10*-GERMS. J. Am. Soc. Mass Spectrom. 24, 1185–1193. http://dx.doi.org/10.1007/s13361-013-0627-8.

Teramoto, K., Sato, H., Sun, L., Torimura, M., Tao, H., Yoshikawa, H., Hotta, Y., Hosoda, A., Tamura, H., 2007. Phylogenetic classification of *Pseudomonas putida* strains by MALDI-MS using ribosomal subunit proteins as biomarkers. Anal. Chem. 79, 8712–8719.

Whittam, T.S., Wolfe, M.L., Wachsmuth, I.K., Orskov, F., Orskov, I., Wilson, R.A., 1993. Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. Infect. Immun. 61, 1619–1629.