C146-E331

SHIMADZU
Excellence in Science

Technical Report

Serotype-Level Bacterial Discrimination Using Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry

Practical Application of the Highly-Accurate Bacterial Discrimination Software "Strain Solution Ver. 2" with the iD*plus*

Teruyo Ojima-Kato[1], Hiroto Tamura[1, 2], Naomi Yamamoto[2], Keisuke Shima[3]

Abstract:

We successfully achieved serotype-level discrimination of enterohemorrhagic *Escherichia coli* using a proteotyping method (the *S10*-GERMS method) with markers consisting of ribosomal proteins detected during bacterial measurements using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS). In the future, this technique could be utilized as a routine method for the discrimination of bacteria at levels finer than the species level, which has thus far not been possible with conventional MALDI-TOF MS fingerprinting-based microbial identification methods.

Keywords: food-poisoning pathogens, MALDI-TOF MS, proteotyping, *S10*-GERMS

1. Introduction

Fast and accurate microbial identification is needed on a routine basis during the manufacturing of food products and pharmaceuticals, as well as in clinical microbiological testing. In particular, the rapid identification of pathogenic microorganisms and microbial contaminants (opportunistic bacteria) has become increasingly important. Conventional microbial identification methods generally include physiological and biochemical tests as well as DNA nucleotide sequence analyses targeting 16S rRNA gene sequences. However, these techniques are associated with the following challenges: 1) they lack rapidity; 2) they are labor-intensive and require expertise; 3) for some bacterial species, detailed identification is impossible beyond the genus or species level; and, importantly, 4) the discrimination of pathogenic microorganisms requires the use of costly reagents (e.g., antisera/antibodies). Thus, methods based on matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) have received wide attention as new techniques for the rapid identification of microorganisms. MALDI-TOF MS is attractive owing to its simplicity and convenience, and the fact that it allows for the rapid processing of multiple samples. Basically, analyses can be carried out by simply mixing microbial samples from a single colony with an extremely small amount of matrix solution (ionization adjuvant). In addition, the required analysis time of this method is less than 1 minute per sample, which allows for processing multiple samples in a short amount of time. Microbial identification methods using MALDI-TOF MS have been spreading rapidly since the late 2000s, mainly in the field of clinical microbiological testing, and were approved by the United States Food and Drug Administration in 2013. The procedure of microbial identification using MALDI-TOF MS consists of acquiring the mass spectrum of a microbial sample and matching it with the mass spectra of various bacterial species that are previously recorded in a database (fingerprinting method). Although the fingerprinting method allows for identification up to the genus and species levels, it does not allow for detailed discrimination of the serotypes of important food poisoning-related pathogens such as *Escherichia coli* and *Salmonella*. Therefore, we developed a method for discrimination of the serotypes of food poisoning-causing bacteria using the proteotyping method*1 *S10-spc-alpha* operon Gene Encoded Ribosomal protein Mass Spectrum (*S10*-GERMS method) [1], which enables discrimination of microorganisms with high accuracy, surpassing that of the fingerprinting method. In addition, a highly-accurate bacterial discrimination software named Strain Solution Ver. 2 was developed, which is compatible with the iD*plus* (Fig. 1). In this communication, the use of MALDI-TOF MS and Strain Solution Ver. 2 for the discrimination of microorganisms to the level of serotypes is introduced, and an example of the application of the method for enterohemorrhagic *Escherichia coli**2 is described [2].

*1 proteotyping method:
In this method, protein components detected by MALDI-TOF MS are used as biomarkers for the discrimination of microorganisms. The biomarkers used for proteotyping are specified in advance on the basis of the nucleotide sequence data of the target gene. Compared to typing methods based on genetic techniques such as the conventional DNA sequencing method, proteotyping using MALDI-TOF MS offers the advantages of being faster, simpler, and easier to use. If the genome sequence of the targeted bacterial species has not yet been decoded, proteotyping can still be carried out by decoding the base sequence of the *S10-spc-alpha* operon, which encodes approximately half of all ribosomal proteins. The *S10*-GERMS method was developed by the Department of Environmental Bioscience, Faculty of Agriculture, Meijo University, and by the Environmental Measurement Technology Group, National Institute of Advanced Industrial Science and Technology.

*2 Enterohemorrhagic *Escherichia coli*:
Nearly 200 different serotypes of *Escherichia coli* (type O antigen) have been reported thus far, but the Shiga toxin-producing serotypes O157, O26, and O111 are the major pathogens responsible for food poisoning, and are therefore the most problematic. However, bacterial identification methods using conventional fingerprinting techniques are limited to the species level, which does not allow for determination of serotypes.

2. The *S10*-GERMS Method and Strain Solution Ver. 2

The *S10*-GERMS method identifies the genes encoding the proteins (biomarkers) detected as mass peaks by MALDI-TOF MS that are specific to serotypes or strains of microorganisms. As a result, a database of the theoretical masses of mass peaks that could potentially be used as biomarkers for the distinction between serotypes or strains can be constructed on the basis of the DNA sequence information of the gene as well as on the basis of the actual measured values. Strain Solution Ver. 2 matches the list of the mass peaks obtained from the sample with the database of theoretical values.

Unlike fingerprinting methods, the components of the biomarkers used in this method are well-defined, and therefore discrimination can be achieved even if the difference consists of only a single amino acid mutation resulting from a single-base mutation. For this reason, this method is also useful for the molecular phylogenetic analysis that follows microbial identification, as well as for the analysis of mixed samples (which will be described later).

1 Research Institute, Meijo University
2 Department of Environmental Bioscience, Faculty of Agriculture, Meijo University
3 Analytical & Measuring Instruments Division, Shimadzu Corporation

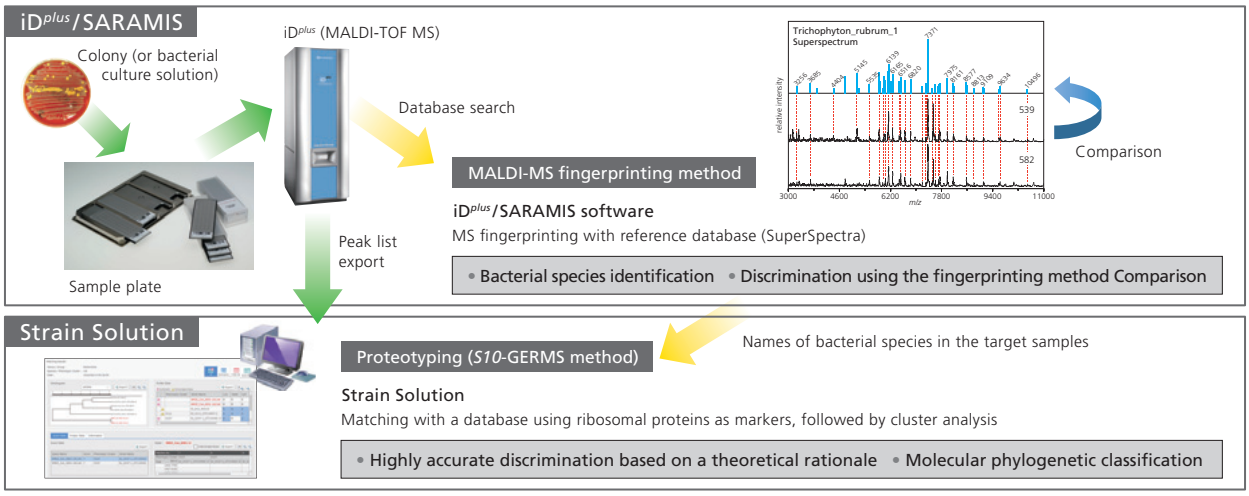


Fig. 1 iD^{plus} and Strain Solution

3. Application of the Method

Here, we introduce the development of a database aimed at distinguishing among the major serotypes of enterohemorrhagic *Escherichia coli*, namely O157, O26, and O111, and provide an example of the analysis using Strain Solution Ver. 2.

3-1. Development of a Database Designed for Use with Strain Solution Ver. 2

First, 50 strains of *Escherichia coli* with various types of O antigens were obtained from culture collections; measurements were carried out with MALDI-TOF MS (iD^{plus}), and biomarker proteins that could potentially be used as indicators for serotype discrimination were selected. Genomic information or experimentally analyzed DNA sequence data were matched with data from actual measurements using MALDI-TOF MS, and a table of the theoretical masses of the selected biomarkers was created (Fig. 2). As a result, we found a number of biomarker candidates that showed mass numbers characteristic to each *Escherichia coli* serotype, and we were able to distinguish 12 groups (groups A–L) among the 50 strains of *Escherichia coli*.

Among these groups, we found that the presence or absence of the acid-stress protein HdeB (*m/z* 9066.2), the ribosomal protein subunits S15 (*m/z* 10166.6 or 10138.6) and L25 (*m/z* 10676.4 or 10694.4), as well as the DNA-binding protein H-NS (*m/z* 15409.4 or 15425.4) were characteristic of serotypes O157, O26, and O111, the major serotypes of enterohemorrhagic *Escherichia coli*. These biomarkers were also useful for distinguishing between the aforementioned serotypes and other *Escherichia coli* strains and serotypes.

3-2. Verification Using Wild-Type Strains

We carried out a verification of the practicality of the developed database using 45 wild-type strains of *Escherichia coli* of various serotypes, which were isolated from food poisoning patients as well as from food samples (Table 1) [3]. The strains were cultured in tryptone soy agar medium and common *Escherichia coli* selective media (desoxycholate agar medium, CT-SMAC medium, Chromagar X-gal medium, and VRBL medium). Specifically, the following steps were carried out for the verification (Fig. 1).

(1) On the basis of the theoretical mass list shown in Fig. 2, the marker peaks of HdeB, S15, L25, and H-NS of *Escherichia coli* strains O157, O26, O111, and K12 were recorded in the Strain Solution Ver. 2 database.

(2) Single colonies (approximately 10^7 cells) of the wild-type strains used in the verification experiments were cultured on agar medium and smeared on metal plates for analysis, and then mixed with 1 μL of matrix solution (20 mg/mL sinapic acid, 50% acetonitrile, 1% trifluoroacetic acid).

(3) Analyses were carried out using the iD^{plus}, and mass spectra were obtained.

(4) Using Strain Solution Ver. 2, the list of mass peaks in the resulting mass spectrum was matched with the marker peaks registered in the database.

Table 1 *Escherichia coli* Strains Used in the Verification Experiments [3]

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.

Theoretical mass database

		Group of mass pattern																	
		A	B	C		D				E	F	G		H		I	J	K	L
Number of strains		14	2	2	9	3	1	1	4	1	1	1	2	5	1	1			
Protein	Coded operon	O157	O157	O111	O26	O121	O128	O152	-	O115	O119	O63	-	K12	-	-	-	-	O150
L23	S10	11200.1	11200.1	11200.1	11200.1	11200.1	11200.1	11200.1	11200.1	11147.1	11200.1	11200.1	11200.1	11200.1	11200.1	11200.1	11200.1	11200.1	11200.1
L24	spc	11186.0	11186.0	11186.0	11186.0	11186.0	11186.0	11186.0	11186.0	11186.0	11186.0	11216.0	11216.1	11186.0	11186.0	11186.0	11186.0	11186.0	11216.1
S14	spc	11450.3	11450.3	11450.3	11450.3	11450.3	11450.3	11450.3	11450.3	11450.3	11450.3	11450.3	11450.3	11450.3	11450.3	11450.3	11464.3	11450.3	11450.3
L15	spc	14967.4	14967.4	14967.4	14967.4	14967.4	14967.4	14967.4	14967.4	14981.4	14945.0	14967.4	14967.4	14981.4	14981.4	14981.4	14967.4	14967.4	14967.4
S11+Me	alpha	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13728.8	13756.8
S15		**10166.6**	10138.6	10138.6	10138.6	10138.6	10138.6	10138.6	10138.6	10138.6	10138.6	10138.6	10138.6	10138.6	10138.6	10137.6	10138.6	10138.6	10138.6
L25		**10676.4**	10694.4	10694.4	10694.4	10694.4	10694.4	10694.4	10694.4	10694.4	10694.4	10693.5	10693.5	10694.4	10694.4	10694.4	10694.4	10693.5	10693.5
HdeB		-	-	9066.2	9066.2	9066.2	9066.2	9066.2	9066.2	9066.2	9066.2	9066.2	9066.2	9066.2	9066.2	9066.2	9066.2	9066.2	9066.2
H-NS		15409.4	15409.4	**15425.4**	**15425.4**	15409.4	15409.4	15409.4	15409.4	15409.4	15409.4	15409.4	15409.4	15409.4	15409.4	15882.0	15409.4	15409.4	15409.4

Biomarker peaks allowing for discrimination among *Escherichia coli* O157, O26, and O111

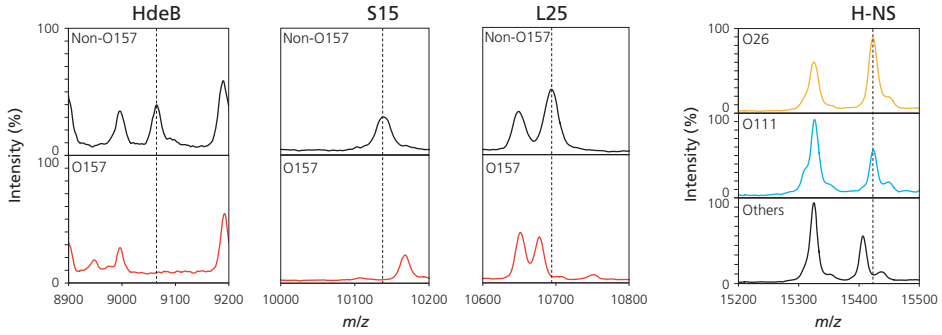


Fig. 2 Construction of the Database Used for Distinguishing Between Serotypes of Enterohemorrhagic *Escherichia coli* on the Basis of the *S10*-GERMS Method.

The results of discrimination using the cluster analysis function of Strain Solution Ver. 2 are shown in Fig. 3. The strains were broadly classified into the following four clusters: serotype O157 (cluster A), other serotypes (cluster B), serotypes O26 and O111 (cluster C), and serotype O157 (cluster D). The percentage rates of correct discrimination of the serotypes O157 (12 strains) and O26/O111 (13 strains) were 92% (11/12 strains) and 100% (13/13 strains), respectively. These findings showed a high degree of concordance with the results of the determination of serotypes using conventional methods. However, the 11 strains of the Shiga toxin-producing serotype O157 were classified with a 100% discrimination rate in the cluster of O157 (D), revealing that *Escherichia coli* O157 KB0341-2 (a non-shigatoxin producing strain), which was classified as "others" (cluster B), was different from the common O157 strains. In addition, O121 KB0747 had no HdeB peak; therefore, it was classified in the same group (group A) as GTC14550, a rare strain within the O157 serotype.

Fig. 3 shows the results of measurements conducted on colonies that were grown on tryptone soy agar medium, but similar results were also obtained from strains grown in various types of selective media. Of course, it is necessary to carry out verifications on a number of other samples as well; however, these results showed that our database and serotype discrimination by proteotyping seems to be effective for most wild-type strains. Another advantage of this method is that it allows for analyses to be conducted for colonies grown on selective culture media that are used in official methods as well as for those used in independent tests. As a result, this method can be applied to a wide variety of samples without having to change the conventional microbial testing process. The database described in this study is available as an option in the software Strain Solution Ver. 2. Further original databases can also be created depending on the purpose.

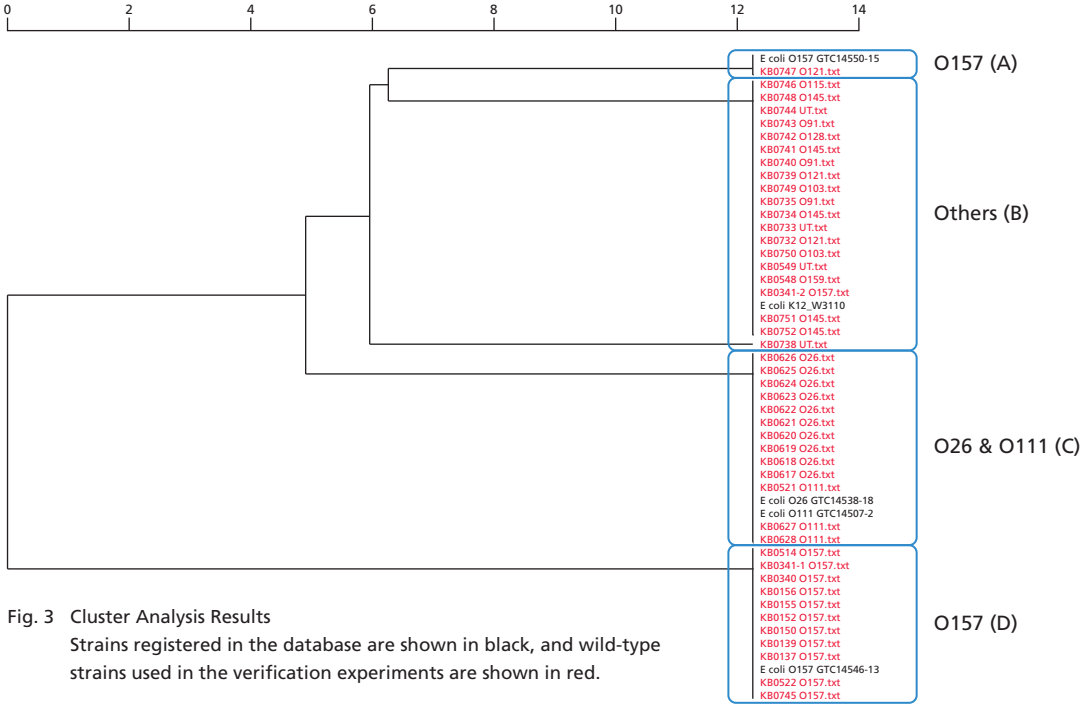
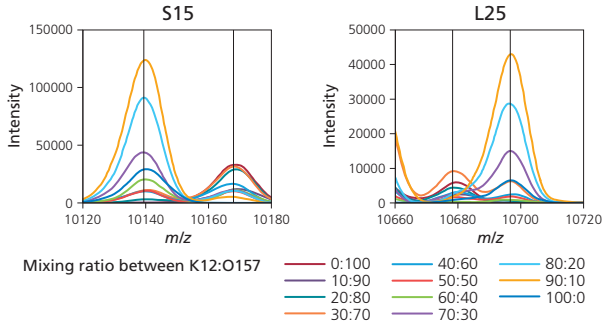


Fig. 3 Cluster Analysis Results
Strains registered in the database are shown in black, and wild-type strains used in the verification experiments are shown in red.

3-3. Mixed Samples

Here, *Escherichia coli* strain K12, which is a common strain used in experiments, was mixed at various ratios with the enterohemorrhagic *Escherichia coli* strain O157, and analyses were carried out using MALDI-TOF MS to verify whether Strain Solution Ver. 2 could identify the mixtures with three biomarkers as indicators. The results showed that when O157 was mixed at ratios of 10% to 80%, the software was able to correctly determine that the samples were mixtures of O157 and other *Escherichia coli* (K12 strains). When the samples were mixed at ratios of 40% to 70%, all three biomarkers were recognized as two mass peaks (Fig. 4) [3].

Biomarker mass peaks



Results from the Strain Solution

Mixing ratio	Hit biomarker			Results of identifications using the Strain Solution
K12:O157	S15	L25	H-NS	
100:0	C	C	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

A: Mass corresponding to the biomarker of the O157 type (*Escherichia coli*).
C: Mass corresponding to the biomarker of other *Escherichia coli*

Fig. 4 Discrimination of Mixed Bacteria

Similar to conventional microbial identification methods, bacteria normally need to be isolated prior to analysis with MALDI-TOF MS; however, if closely related bacterial strains are mixed together, they cannot be distinguished from each other through the standard fingerprinting method. Therefore, such functions of Strain Solution Ver. 2 are useful for monitoring human error when dealing with several closely related bacterial strains. This is beneficial for the management of the quality of starter bacterial strains used in the food industry, as well as culture collections for the management of their microbial collections.

4. Conclusion and Future Prospects

Due to the rapidity and simplicity of MALDI-TOF MS, its use and applications are rapidly expanding in various fields, including for microbial control such as in clinical settings, the pharmaceutical industry, and the food industry. Several studies have reported the use of MALDI-TOF MS for direct measurements of bacteria from blood culture media and food samples in recent years, and simpler analyses with higher accuracy are needed. The proteotyping technique using Strain Solution Ver. 2 based on the *S10*-GERMS method allows for distinguishing the small changes in the mass of selected biomarkers with high sensitivity, which has thus far not been considered important. Therefore, this method can potentially be utilized in a wide range of applications, not only in the discrimination of serotypes or strains of single microorganisms but also in the detailed analysis and diagnosis of the intestinal flora or bacterial mixtures that do not allow for single-strain isolation.

Acknowledgments

This work was financially supported by the Aichi Science and Technology Foundation (Japan) as a part of the "Technological Development Project for Food Safety and Security" under Knowledge Hub Aichi. Verification of wild-type strains was conducted with the cooperation of Kobe City's Environmental Health Research Institute.

References

[1] Yudai Hotta, Kanae Teramoto, Hiroaki Sato, Hiromichi Yoshikawa, Akifumi Hosoda, and Hiroto Tamura: Classification of Genus Pseudomonas by MALDI-TOF MS Based on Ribosomal Protein Coding in *S10-spc-alpha* Operon at Strain Level. *J. Proteome Res.*, 2010, 9, 6722–6728.

[2] Teruyo Ojima-Kato, Naomi Yamamoto, Mayumi Suzuki, Tomohiro Fukunaga, and Hiroto Tamura: 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.

[3] Teruyo Ojima-Kato, Naomi Yamamoto, Yoshio Iijima, and Hiroto Tamura: 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. *J. Microbiol. Methods*, 2015, 119, 233–238.

Strain Solution software was created on the basis of the results obtained from the Food Safety & Security Technology Development Project (group leader: Professor Hiroto Tamura of the Laboratory of Environmental Microbiology, Faculty of Agriculture, Meijo University), a "Knowledge Hub Aichi" core research project of Aichi Prefecture, which is in-turn based on ideas from the Meijo University and the National Institute of Advanced Industrial Science and Technology.

First Edition: December, 2016

SHIMADZU

Shimadzu Corporation

www.shimadzu.com/an/

Printed in Japan 3655-10603-10ANS