C146-E244

SHIMADZU
Excellence in Science

Technical Report

# Characterization of the *Lactobacillus Casei* Group Based on Profiling of Ribosomal Proteins Coded in *S10*-spc-alpha Operons as Observed by MALDI-TOF MS

Keisuke Shima[1], Hiroaki Sato[2], Moriya Ohkuma[3], Hiroto Tamura[4]

## Abstract:

The taxonomy of the members of the *Lactobacillus casei* group is complicated due to the genetic similarities within the group and controversial nomenclature status. We were able to conduct quick and accurate discrimination of the *L. casei* group using as markers ribosomal proteins detected during measurement of bacteria by MALDI-TOF MS. This technique can be considered as a molecular phylogenetic method based on theoretical principles, and may therefore be useful for such applications as management of bacterial strains found in humans or animals, which cannot be accomplished by existing bacterial strain discrimination using MALDI-TOF MS fingerprinting.

**Keywords: *lactobacillus*, molecular phylogenetic analysis, ribosomal protein, MALDI-TOF MS, *S10*-GERMS**

## 1. Introduction

*Lactobacillus* is widely distributed in nature, inhabiting the bodies of animals, and residing in foods and on the surface of plants. While it is representative of a beneficial type of bacterium that is used in the processing and manufacturing of fermented foods including yogurt, cheese and pickles, etc., *lactobacillus* is also known to play a role in the deterioration of foods and beverages.

Among lactobacilli, *Lactobacillus casei* and related species (*L. casei* group) are typical beneficial bacteria that are widely used in such applications as starters for *lactobacillus* beverages. However, despite the wide applicability of the *L. casei* group based on its usefulness, its taxonomic position assignment and nomenclature are not clearly established.

Because the strains within the *L. casei* group are closely related genetically to one another, identification based on the 16S rRNA gene sequence is difficult. In addition to the technique based on the sequence of housekeeping genes, techniques such as MLST (Multilocus sequence typing) are being used to discriminate within the group. However, there have been problems associated with these genetic techniques, not the least of which are the time and effort required to conduct DNA extraction and amplification, in addition to DNA sequencing.

MALDI-TOF MS (matrix-assisted laser desorption ionization time-of-flight mass spectrometry), a technique that has attracted attention in recent years for the identification and classification of microorganisms, may be a useful method for addressing the shortcomings of the classification and discrimination method based on genetic techniques, thanks to its speed and convenience. However, use of the fingerprinting method* with the currently available MALDI-based microbial identification systems poses some challenges. For example, its theoretical foundation and reliability in discriminating among closely related species, such as the L. casei group, makes its application to molecular phylogenetic classification substantively difficult.

Utilizing the fact that about half of the peaks observed in MALDI-TOF MS measurement of the bacterial cell body are ribosomal proteins, an association is made between the mass of the peak observed using MALDI-TOF MS and the predicted calculated mass from the amino acid sequence translated from the ribosomal protein gene base sequence information. By ascribing the type of protein that is derived from that peak, the species associated with that information can be identified. Furthermore, by selecting marker peaks attributed to the protein of origin based on this method, even in species closely related to each other such as the *L. casei* group, highly reliable molecular phylogenetics based on theoretical grounds becomes possible.

However, to conduct molecular phylogenetics based on this approach, decryption of the genetic sequence of the reference strain was necessary to obtain the ribosomal protein genetic information of the bacterial species of interest. Since almost none of the genomic sequence of the *L. casei* group type strain had previously been decoded, direct application of this technique was difficult. Focusing on the *S10-spc-alpha* operon, of which approximately half of the ribosomal protein has previously been encoded, MALDI-TOF MS was used to decipher the nucleotide sequence of this operon.

In this Technical Report, we introduce the results of the cited Reference[1] related to the classification of the *L. casei* group using the *S10*-GERMS Method (*S10-spc-alpha* operon Gene Encoded Ribosomal Protein Mass Spectrum) developed by Japan's National Institute of Advanced Industrial Science and Technology and the Laboratory of Environmental Microbiology, Faculty of Agriculture, Meijo

1 Global Application Development Center, Analytical & Measuring Instruments Division, Shimadzu Corporation
2 Research Institute for Environmental Management Technology, National Institute of Advanced Industrial Science and Technology
3 Microbe Division / Japan Collection of Microorganisms, RIKEN BioResource Center
4 Laboratory of Environmental Microbiology, Faculty of Agriculture, Meijo University

University. Also introduced are the results obtained using Strain Solution, a high-precision bacterial identification software application for AXIMA MALDI-TOF MS microbial identification systems that was developed based on the *S10*-GERMS Method.

* A method of identification in which unique mass spectral information from a known microorganism is entered into a database beforehand, and identification is conducted by matching information from the mass spectrum of unknown microorganism samples to entries in the database.

Table 1 List of Strains Tested

| Species name | Strains | Ribotype[2] |
|---|---|---|
| *L. casei* | NBRC 15883$^T$ (=JCM 1134$^T$) | C5 |
| | JCM 8129 | C5 |
| | JCM 11302 | C4 |
| *L. parasacei* subsp. *paracasei* | JCM 1053 | C1 |
| | JCM 1109 | C1 |
| | JCM 1111 | C1 |
| | JCM 1133 | C1 |
| | JCM 1161 | C1 |
| | JCM 1163 | C1 |
| | JCM 1181 | C1 |
| | JCM 1556 | C1 |
| | JCM 2123 | |
| | JCM 2769 | C1 |
| | JCM 8130$^T$ | C1 |
| | JCM 8131 | C1 |
| | JCM 8132 | C1 |
| | JCM 8133 | C1 |
| *L. parasacei* subsp. *tolerans* | JCM 1171$^T$ | C3 |
| *L. rhamnosus* | JCM 1136$^T$ | C2 |
| | JCM 1165 | C2 |
| | JCM 1553 | C2 |
| | JCM 1561 | C2 |
| | JCM 1563 | C2 |
| | JCM 2771 | C2 |
| | JCM 2772 | C2 |
| | JCM 8134 | C2 |
| | JCM 8135 | C2 |
| | JCM 8136 | C2 |

# 2. Method

## 2-1. Tested Microorganism Strains and Culture Conditions

The *L. casei* group bacterial strains were obtained from the Microbe Division of the Japan Collection of Microorganisms, RIKEN BioResource Center (JCM) and the Biological Resource Center, National Institute of Technology and Evaluation (NBRC). (Table 1)

The culture was prepared in accordance with the conditions specified by the vendors.

## 2-2. Database Construction for Ribosomal Protein and DNA Sequencing of *S10-spc-alpha* Operon Strain

The base sequences of the 4 reference strains of the *L. casei* group (*L. casei* NBRC 15883$^T$, *L. paracasei* subsp. *paracasei* JCM 8130$^T$, *L. paracasei* subsp. *tolerans* JCM 1171$^T$, *L. rhamnosus* JCM 1136$^T$) and that of the *L. casei* JCM 11302 *S10-spc-alpha* operon were determined by DNA sequencing of the PCR products serving as genomic DNA templates prepared from test bacteria, using the primer sets designed based on the nucleotide sequence of the genome sequencing strain. The ORF ribosomal protein subunit was determined based on the genome sequencing strain information from the determined base sequence, after which it was replaced by the amino acid sequence. The theoretical molecular weights of their amino acid sequences were calculated taking into consideration post-translational modifications and cleavage of the N-terminal methionine residues. Also, since what is actually observed using MALDI-TOF MS are molecules to which protons have been added, we built a list of the theoretical masses of the ribosomal proteins taking into account those protons. As a final step, we validated the prepared list of theoretical masses by comparing it with the mass spectra of the reference strains that were measured by MALDI-TOF MS.

This prepared list was registered in advance in the Strain Solution high-precision bacterial identification software installed in the AXIMA MALDI-TOF MS Microbial Identification System, and a ribosomal protein database of the five strains of the *L. casei* group described above was constructed. In addition, data associated with strains of ribosomal proteins other than strains of the JCM8129 were also registered beforehand in the database constructed in Strain Solution, based on actual measured values using MALDI-TOF MS and the reference strain data.

## 2-3. Measurement by MALDI-TOF MS

Measurement mode: Linear positive

Matrix solution: Sinapic acid 10 mg/mL (50% acetonitrile, 1% TFA)

Sample preparation: Bacteria bodies suspended in TMA -I buffer solution (10 mM Tris-HCl (pH 7.8), 30 mM $NH_4Cl$, 10 mM $MgCl_2$, and 6 mM 2-mercaptoethanol) were disrupted using beads to prepare a cell lysate. The prepared suspension of disrupted cells was mixed with matrix solution at a 1:4 ratio, and approximately 2 µL of the suspension was mounted on a sample plate and allowed to dry naturally. The prepared sample was then subjected to measurement.

## 2-4. Matching of Ribosomal Protein Peaks Using Strain Solution and Molecular Phylogenetic Classification Using External Software

Matching of the *L. casei* JCM 8129 peak list obtained by MALDI-TOF MS measurement was conducted using Strain Solution, and analysis was conducted to determine the *L. casei* group species to which JCM 8129 belongs.

Matching was conducted on the *L. casei* group ribosomal protein database that had been registered in the Strain Solution beforehand.

Regarding the matching results, when a peak of the *L. casei* group that had been registered in the database matched a JCM 8129 peak, a "1" was assigned, and when it did not match, a "0" was assigned, thereby creating a binary (1/0) format profile. By applying the cluster analysis algorithm UPGMA (Unweighted Pair Group Method with Arithmetic Mean) on this profile using BioNumerics (Applied Maths, Japan representative: Infocom Co.), we were able to analyze which of the *L. casei* group reference strains to which JCM 8129 belonged.

Table 2 List of Ribosomal Proteins of *L. casei* Group Reference Strains and *L. casei* JCM 11302

| Protein name | Coded operon | Theoretical masses as $[M+H]^+$ | | | | |
|---|---|---|---|---|---|---|
| | | *L. casei* NBRC 15883$^T$ | *L. paracasei* subsp. *paracasei* JCM 8130$^T$ | *L. paracasei* subsp. *tolerans* JCM 1171$^T$ | *L. rhamnosus* JCM 1136$^T$ | *L. casei* JCM 11302 |
| L22 | *S10* | 12599.5 | 12599.5 | 12599.5 | 12599.5 | 12599.5 |
| L23 | *S10* | 11564.4 | 11561.4 | 11589.5 | 11533.4 | 11564.4 |
| L29 | *S10* | 7578.9 | 7479.7 | 7479.7 | 7578.9 | 7578.9 |
| S17 | *S10* | 9985.5 | 10071.7 | 10071.7 | 9985.5 | 9985.5 |
| S19 | *S10* | 10447.0 | 10447.0 | 10447.0 | 10447.0 | 10447.0 |
| L14 | *spc* | 13031.2 | 13036.2 | 13036.2 | 13036.2 | 13031.2 |
| L18 | *spc* | 13015.8 | 13001.8 | 13001.8 | 13015.8 | 13015.8 |
| L24 | *spc* | 11231.1 | 11231.1 | 11203.0 | 11247.1 | 11231.1 |
| L30 | *spc* | 6661.8 | 6618.8 | 6618.8 | 6661.8 | 6661.8 |
| S08 | *spc* | 14676.1 | 14644.0 | 14644.0 | 14644.0 | 14676.1 |
| S14 | *spc* | 6992.3 | 6992.3 | 7006.4 | 7006.4 | 6992.3 |
| L36 | *alpha* | 4449.5 | 4449.5 | 4449.5 | 4449.5 | 4449.5 |
| S11 | *alpha* | 13641.6 | 13641.6 | 13641.6 | 13641.6 | 13641.6 |
| S13 | *alpha* | 13426.6 | 13360.5 | 13360.5 | 13426.6 | 13412.6 |

# 3. Results

## 3-1. *L. casei* Group Ribosomal Protein Database

Table 2 shows the list of base sequences of the *S10-spc-alpha* operon and the ribosomal protein theoretical masses for the *L. casei* group reference strains and *L. casei* JCM 11302, which was created based on the measured values by MALDI-TOF MS.

There are four ribosomal proteins (L36, S19, L22, S11) with the same theoretical mass in each strain, suggesting the possibility that these peaks can be used as group-specific markers to distinguish among those that do and do not belong to the *L. casei* group.

Since the theoretical masses of other ribosomal proteins are shared among several strains or are strain-specific, it is also suggested from this table that these peaks can be used as markers to characterize each reference strain.

## 3-2. Data Obtained from MALDI-TOF MS Measurement

As can be inferred from the similarity of the genetic characteristics among the groups, the patterns of the mass spectra within the *L. casei* group are similar to one another. In this situation, it would probably be difficult using existing MALDI-TOF MS fingerprinting to distinguish, with good repeatability and based on theoretical principle, between *L. casei* NBRC 15883$^T$ and *L. casei* JCM 11302, or between *L. paracasei* subsp. *paracasei* JCM 8130$^T$ and *L. paracasei* subsp. *tolerans* JCM 1171$^T$, in particular.

However, even among the many peaks that match one another, the masses of peaks originating from the L23 and L24 ribosomal protein large subunits and from the S13 small subunit vary according to the reference strain, as suggested by Table 2 and as shown in the MALDI-TOF MS measurement data in Fig. 1. Therefore, it may be possible to theoretically determine which type of (reference) strain the *L. casei* group strain resembles.

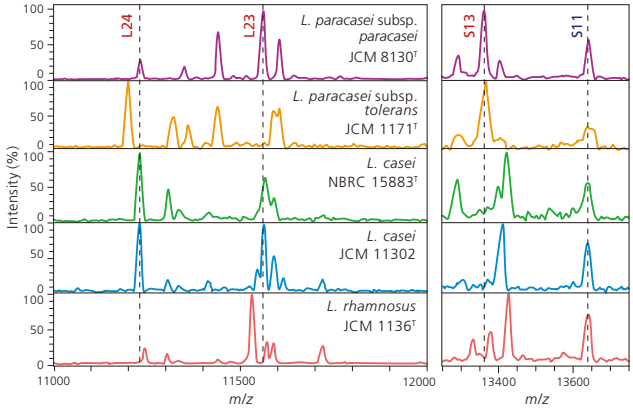


Fig. 1 Mass Spectra of *L. casei* Group Type Strains and *L. casei* JCM 11302 (Magnified)

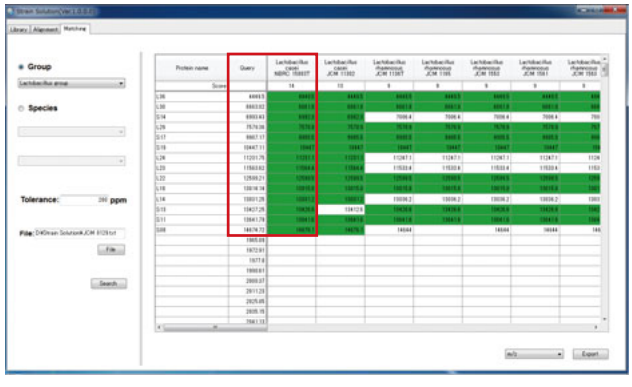

Fig. 2 Matching Results Obtained Using Strain Solution

## 3-3. Matching Using Strain Solution

Fig. 2 shows the matching results generated using Strain Solution with respect to the JCM 8129 peak list obtained by MALDI-TOF MS. The matching results indicate that *L. casei* NBRC 15883$^{T}$ (=JCM 1134$^{T}$) was the strain with the most number of matching peak masses. As for JCM 8129 and NBRC 15883$^{T}$, all 14 of their ribosomal protein marker peaks matched. From this result, it was presumed that JCM 8129 is the most closely related to NBRC 15883$^{T}$ of the registered *L. casei* group. It should be noted that these results matched the Reference[2] obtained using the conventional ribotyping genetic technique.

## 3-4. Molecular Phylogenetic Classification Using External Software

We conducted cluster analysis based on the binary (1/0) format profile obtained from matching, and the results of molecular phylogenetic classification are shown in Fig. 3.

Each of the reference strains was formed into a cluster of its own (85% cut-off value). Regarding strains other than the type strain, they were all assigned to the same cluster of the type strain corresponding to the species-subspecies name. This result was about the same as that obtained in the classification results of Reference[2] using past genetic techniques.

In addition to those indicated in the above results, it should be noted that redefinition of the species name is required for some strains of the *L. casei* group as cited in the Reference[1] report. There is also a case in which a mistaken species name is used for commercially available lactic acid bacteria, as cited in the Reference[3] report. The proposed method presented here is expected to become a powerful tool in helping to resolve these problems by providing rapid and accurate discrimination of the *L. casei* group.

## 4. Conclusion

- The *S10*-GERMS method, a technique that uses a combination of the base sequence information of the ribosomal protein gene and MALDI-TOF MS, permitted quick and accurate discrimination within the *L. casei* group at the subspecies level, which up to now has been difficult using MALDI-TOF MS fingerprinting.
- The Strain Solution software that was eventually developed based on the principle of the *S10*-GERMS method may be applicable to various fields requiring fast and accurate identification, such as management of bacterial strains found in humans or animals.

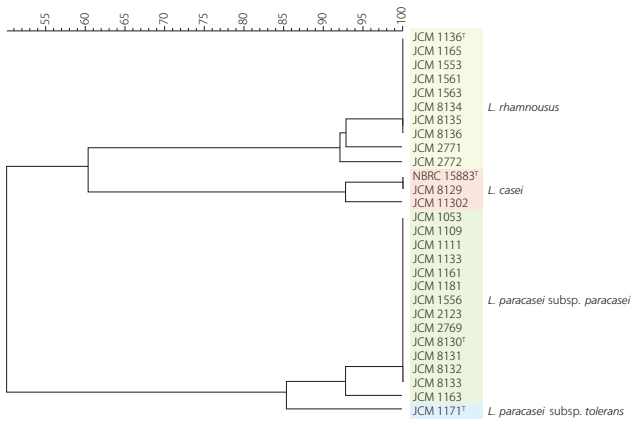


Fig. 3 *L. casei* Group Cluster Analysis (Algorithm: UPGMA) Results Based on Ribosomal Protein Profiling by *S10*-GERMS


References:

1 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.

2 Ryu, C. S., Czajka, J. W., Sakamoto, M., Benno, Y. (2001) Characterization of the *Lactobacillus casei* group and the *Lactobacillus acidophilus* group by automated ribotyping. *Microbiol. Immunol*. 45, 271–275.

3 Huys, G., Vancanneyt, M., D'Haene, K., Vankerckhoven, V., Goossens, H., Swings, J. (2006) Accuracy of species identity of commercial bacterial cultures intended for probiotic or nutritional use. *Res. Microbiol*. 157, 803–810.



The Strain Solution software was created based on results obtained from an Aichi prefectural key "Knowledge Hub" research project on developing technologies for ensuring food safety and security (group leader: Professor Hiroto Tamura of the Department of Environmental Bioscience, Faculty of Agriculture, Meijo University), based on ideas from Meijo University and the National Institute of Advanced Industrial Science and Technology (AIST).



First Edition: September, 2013



Shimadzu Corporation

**www.shimadzu.com/an/**





Printed in Japan 3655-09301-10ANS