

RESEARCH ARTICLE

# Rapid Classification and Identification of *Microcystis aeruginosa* Strains Using MALDI–TOF MS and Polygenetic Analysis

Li-Wei Sun<sup>1\*</sup>, Wen-Jing Jiang<sup>1</sup>, Hiroaki Sato<sup>2\*</sup>, Masanobu Kawachi<sup>3</sup>, Xi-Wu Lu<sup>1</sup>

**1** School of Energy & Environment, Southeast University, Nanjing, Jiangsu, China, **2** Environmental Measurement Technology Group, Environmental Management Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan, **3** Biodiversity Resource Conservation Section, Center for Environmental Biology and Ecosystem Studies, National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan

\* [liwei-sun@seu.edu.cn](mailto:liwei-sun@seu.edu.cn) (LWS); [sato-hiroaki@aist.go.jp](mailto:sato-hiroaki@aist.go.jp) (HS)



**OPEN ACCESS**

**Citation:** Sun L-W, Jiang W-J, Sato H, Kawachi M, Lu X-W (2016) Rapid Classification and Identification of *Microcystis aeruginosa* Strains Using MALDI–TOF MS and Polygenetic Analysis. PLoS ONE 11(5): e0156275. doi:10.1371/journal.pone.0156275

**Editor:** Brett Neilan, University of New South Wales, AUSTRALIA

**Received:** February 29, 2016

**Accepted:** May 11, 2016

**Published:** May 26, 2016

**Copyright:** © 2016 Sun et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** The present study was supported by the Jiangsu Environmental Monitoring Foundation, China (Project No.1411), to LWS.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

Matrix-assisted laser desorption–ionization–time-of-flight mass spectrometry (MALDI–TOF MS) was used to establish a rapid, simple, and accurate method to differentiate among strains of *Microcystis aeruginosa*, one of the most prevalent types of bloom-forming cyanobacteria. *M. aeruginosa* NIES-843, for which a complete genome has been sequenced, was used to characterize ribosomal proteins as biomarkers and to optimize conditions for observing ribosomal proteins as major peaks in a given mass spectrum. Thirty-one of 52 ribosomal subunit proteins were detected and identified along the mass spectrum. Fifty-five strains of *M. aeruginosa* from different habitats were analyzed using MALDI–TOF MS; among these samples, different ribosomal protein types were observed. A polygenetic analysis was performed using an unweighted pair-group method with arithmetic means and different ribosomal protein types to classify the strains into five major clades. Two clades primarily contained toxic strains, and the other three clades contained exclusively non-toxic strains. This is the first study to differentiate cyanobacterial strains using MALDI–TOF MS.

## Introduction

Toxic cyanobacterial blooms have been identified in fresh and brackish waterbodies for over 100 years [1]. They are known to cause the deaths of wild and domestic animals worldwide [2] and threaten the health of human beings who use waterbodies with such blooms for recreation or drinking [3].

*Microcystis aeruginosa* is one of the most predominant cyanobacterium found in freshwater bodies. Some strains produce hepatotoxins called microcystins (MCs), which have been suspected of causing human hepatocellular carcinoma in China [3]. The World Health Organization (WHO) has set the drinking water standard for MC-LR at 1 µg/L [4]. MCs exist within cyanobacteria and are released after cyanobacterial cells lyse. In addition, their distribution is highly correlated with the distribution of cyanobacteria. MC concentrations shift with the

movement of cyanobacteria in the waterbodies where they occur, thus monitoring cyanobacteria indirectly allows researchers to monitor MCs.

Cell counting [5], pigment analysis [6, 7, 8], and 16S rRNA gene-based molecular techniques [9, 10] have been used to detect and monitor cyanobacteria in samples taken from naturally occurring sources. The disadvantages of these methods are that they are time-consuming and require a skilled expert to perform the analysis, thus limiting their application in routine monitoring. Therefore, more rapid, easy to use, and reliable monitoring methods are required.

Both toxic and non-toxic cyanobacterial strains coexist in the natural environment, and the uncertain distribution of both types of strains prevents the accurate diagnosis of blooms. As a result, the toxigenic potential of cyanobacterial populations has gone unidentified, and corresponding control measures have not been put in place.

The toxicity of cyanobacterial strains cannot be distinguished by microscope. Instead, several molecular typing methods have been established to distinguish toxic and non-toxic strains from one another. These methods include random amplified polymorphic DNA fingerprinting [11], 16S rDNA analysis [8, 10, 12], 16S–23S rDNA internal transcribed spacer analysis [13, 14], analysis of a segment of the phycocyanin operon *cpcBA* intergenic spacer [10, 11], and MC biosynthesis gene *mcy* analysis [10, 15, 16]. However, the results of these analyses have indicated that the toxicity of different strains does not always coincide with that strain's gene type, even when the MC genes are employed as markers (i.e., strains of one *mcy* genotype range from toxic to non-toxic).

Rantala *et al.* [17] found that microcystin synthetase genes were present in the last common ancestor of a large number of cyanobacteria; however, the ability to produce the toxin has been lost in many of the more derived cyanobacteria lineages. Therefore, the toxicity of any given strain must be determined by a method with a finer resolution than the microcystin synthetase gene level.

Tanabe *et al.* [18] developed the multilocus sequence typing (MLST) method based on seven selected housekeeping loci that could distinguish toxic from the non-toxic strains at the genetic level. However, their method is extremely complex, labor intensive, and time-consuming.

In this study, we propose a new approach to this process by using matrix-assisted laser desorption ionization (MALDI) time-of-flight (TOF) mass spectrometry (MS) to create a rapid, high resolution method for identifying and distinguishing toxic *M. aeruginosa* strains from non-toxic strains.

Several reports have demonstrated the feasibility of using MALDI-TOF-MS to identify microorganisms [19–26]. In whole-cell MALDI-TOF MS, characteristic “fingerprint” spectra are obtained from whole (intact) cells, eliminating the biomarker pre-fractionation, digestion, separation, and cleanup steps. The procedure is quick and requires a minimal amount of biological material (sub-colony amounts); therefore, it is suitable for high-throughput routine analysis and has great potential for application in clinical microbiology and environmental monitoring.

The protein biomarkers observed in the mass spectrum are typically highly expressed proteins with housekeeping functions, such as ribosomal or nucleic acid-binding proteins [19, 27, 28]. These proteins are highly conserved in bacteria, which means they may have universal application. Pineda *et al.* [27] were first to propose the use of ribosomal proteins as biomarkers in the identification of bacteria using MALDI-TOF MS; and in our previous studies, we characterized 42 ribosomal proteins from a typical lactic acid bacterium, *Lactobacillus plantarum*, using both two-dimensional gel electrophoresis confirmation and MALDI-MS determination [29]. The ribosomal proteins within *L. delbrueckii* subsp. *bulgaricus* and *Streptococcus thermophilus* have also been characterized through a comparison of MALDI-TOF MS observation

data and available public protein sequence databases [30]; furthermore, the observed masses of some ribosomal subunit proteins have been found to vary within the species, indicating the possibility of classifying bacteria at the strain level. Finally, polygenetic classification of *Pseudomonas putida* strains has occurred using MALDI-TOF MS with ribosomal proteins as biomarkers [31].

In this study, an optimized pre-treatment method to observe ribosomal proteins as major peaks in the mass spectrum and to characterize ribosomal proteins as biomarkers was tested. Fifty-five strains of *M. aeruginosa* from different habitats were then assembled, and their different ribosomal protein types were analyzed using MALDI-TOF MS. A polygenetic analysis using the ribosomal proteins from the 55 *M. aeruginosa* strains revealed high genetic diversity among the strains. Finally, these strains were further assembled into five major clades based on toxicity.

## Materials and Methods

### *M. aeruginosa* strains

Information about the 55 *M. aeruginosa* strains used in this study is listed in [S1 Table](#). The strains were provided by the Microbial Culture Collection at the National Institute for Environmental Studies (MCC-NIES, Tsukuba, Japan). Most of the strains were isolated from freshwater environments in Japan, but the group also included strains (one each) from each of the following countries: Germany, the United Kingdom, Thailand, and Nepal. All but one of the strains (NIES-478) was a clone. Cultures were grown in an MA medium [32] at 20°C for 4 weeks under a 12:12-h light:dark cycle with a photon density of  $15 \mu\text{mol m}^{-2} \text{s}^{-1}$ . Of the sample strains, the genome-sequenced strain NIES-843 was used to characterize the ribosomal proteins and optimize the sample preparation method.

### Preparation of intact cells

First, because *M. aeruginosa* cells have gas vesicles for buoyancy, a mild sonication treatment was applied. The cells were centrifuged at 9,100 *g* for 10 min in a 15-mL tube, and then the precipitate was transferred to a 1.5-mL tube and centrifuged at 20,400 *g* for 10 min. The resulting pellet was re-suspended with 10  $\mu\text{L}$  of a 50% acetonitrile containing 1% trifluoroacetic acid solution and analyzed using MALDI-TOF MS.

### Preparation of cell lysates

The cell pellet (wet weight ranged from 20–40 mg and was measured during the last preparation step) was suspended in a 160  $\mu\text{L}$  of TMA-1 buffer [10 mM Tris-HCl (pH 7.8), 30 mM  $\text{NH}_4\text{Cl}$ , 10 mM  $\text{MgCl}_2$ , and 6 mM 2-mercaptoethanol], transferred to a tube containing 160 mg zirconia silica beads, and ground at 3,000 rpm with a Mini Bead-Beater (Biospec products) for 1 min. The beads and cell debris were removed by centrifuging the mixture at 5,800 *g* for 25 min, and a portion of the resulting cell lysate was analyzed using MALDI-TOF MS.

### Preparation of the concentrated ribosomal fraction

The cell lysate was adjusted to 250  $\mu\text{L}$  with TMA-1 buffer and then centrifuged using an Allegra 64R (Beckman Coulter, Brea, CA, USA) at ca. 64,000 *g* for 4 hours. The concentrated ribosomal fraction was collected in the precipitate. The precipitate was re-suspended with 10  $\mu\text{L}$  of 50% acetonitrile containing 1% trifluoroacetic acid solution and analyzed using MALDI-TOF MS.

## MALDI-TOF MS

The sample preparation, apparatus, and MALDI-TOF MS data acquisition techniques used in this study were similar to those described in our previous papers [29, 30]. Each sample solution was mixed with a 20 mg mL<sup>-1</sup> of sinapinic acid matrix solution in 50% acetonitrile containing 1% trifluoroacetic acid. Approximately 2 µL of the sample matrix mixture was placed onto the MALDI target and dried in open air. All MALDI mass spectra in the range  $m/z$  2000–40000 were recorded using the positive linear mode by averaging 500 individual laser shots from an AXIMA CFR plus time-of-flight mass spectrometer (Shimadzu/Kratos, Kyoto, Japan) equipped with a pulsed N<sub>2</sub> laser ( $\lambda$ , 337 nm; pulse width, 3 ns; frequency, 10 Hz).

## Cluster Analysis

The mass-matching profiles for each sample strain were processed to build a dendrogram using the unweighted pair group method with arithmetic mean (UPGMA) cluster analysis dendrogram and a categorical coefficient. Analyses were conducted using BioNumerics software, version 3.5 (Applied Maths, Kortrijk, Belgium).

## Results

### Characterization of the ribosomal subunit proteins of *M. aeruginosa* NIES-843

First, a list of the calculated molecular weights corresponding to the ribosomal subunit proteins of the *M. aeruginosa* strain NIES-843 was prepared (Table 1). The complete NIES-843 genome had already been sequenced [33]. The ribosomal protein amino acid sequences were retrieved from a public protein database (UniProtKB: <http://www.uniprot.org/>), and then the theoretical molecular weight of each ribosomal protein was calculated using the Compute pI/MW tool ([http://www.expasy.org/compute\\_pi/](http://www.expasy.org/compute_pi/)). The most common post-translation modification rule “N-end rules” [34] was used to determine the possible protonated versions of each protein in the form  $[M+H]^+$  (Table 1).

Fig 1 shows the mass spectrum of the NIES-843 strain from  $m/z$  4000–16000. The  $m/z$  values of the detected peaks were compared to the calculated masses  $[M+H]^+$  of the ribosomal proteins. Thirty-one of the 52 ribosomal proteins were identified in the mass spectrum. Those located between  $m/z$  4000–16000 are labeled in Fig 1. Table 1 summarizes all the assigned ribosomal proteins from *M. aeruginosa* NIES-843, along with their calculated masses and possible post-translation modifications.

### Comparison of the samples' mass spectra values using three different pre-treatment methods

Fig 2 illustrates the mass spectra of samples prepared using three different pre-treatment methods. For the intact cell mass spectrum, many peaks could be observed; but these peaks could not be attributed to ribosomal proteins (Fig 2a). For the cell lysate pre-treatment, the lysate solution was green and slightly viscous, which is primarily attributable to the pigment (chlorophyll) and a mucilage composed of polysaccharides. As a result, the observed peak intensities of the mass spectrum for this pre-treatment exhibited a reduced signal-to-noise ratio, and no ribosomal proteins were observed (Fig 2b). Because ribosomal proteins should exist in the cell lysates, high-speed centrifugation was performed to collect the ribosomal fraction of the cell lysate. The mass spectrum observed for the collected ribosome fraction showed clear ribosomal protein peaks, as seen in Fig 2c. Thirty-one ribosomal proteins were identified in this mass spectrum (Table 1).

**Table 1. Assigned ribosomal subunit proteins of *M. aeruginosa* NIES 843.**

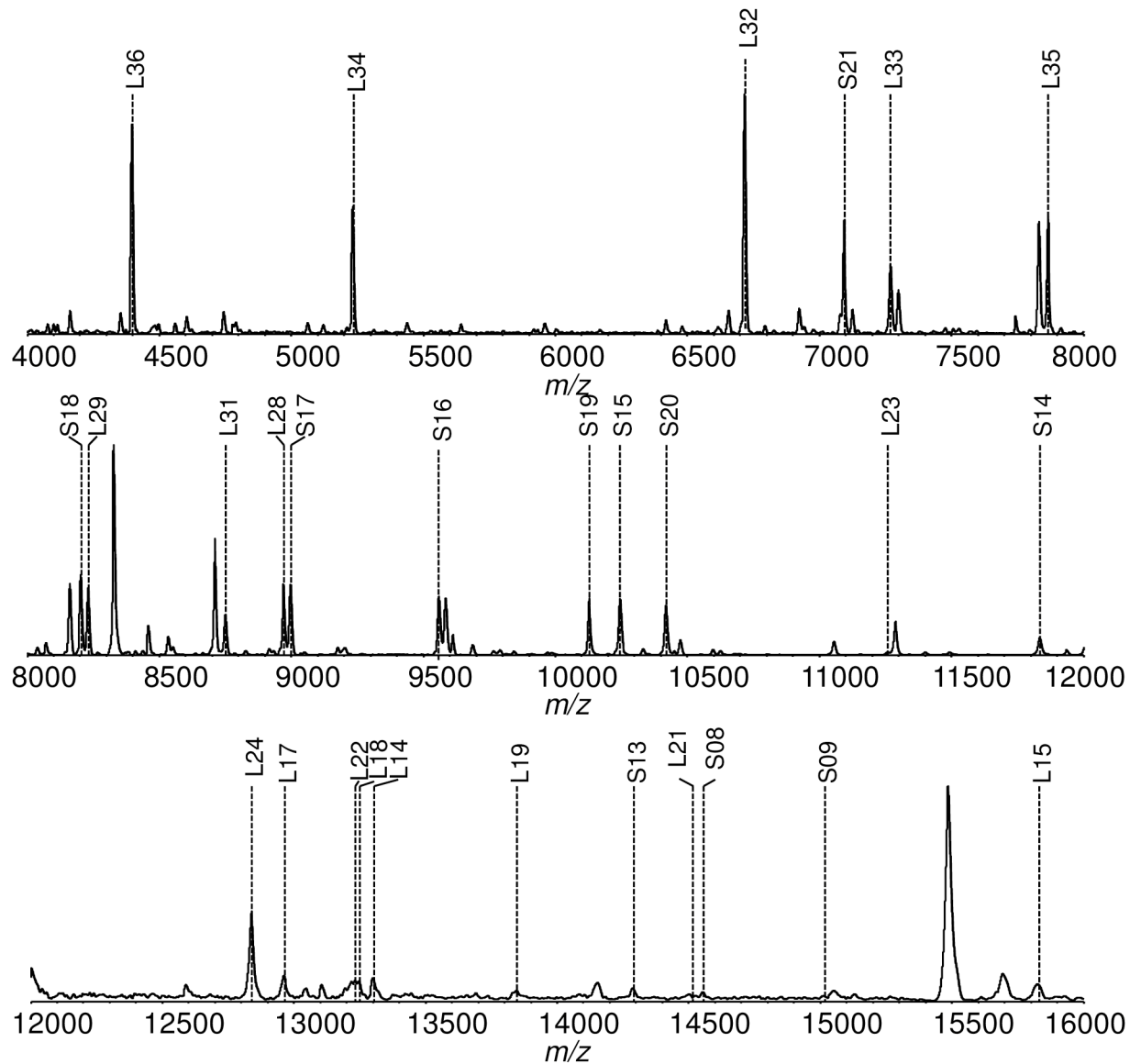
protein name	mass of [M+H]	post-translational modification	second amino acid residue	UniProtKB/Swiss-Prot: access number
<b>Small Subunit Proteins</b>				
S08	14557.9	-Met	S	B0JHZ0.1
S09	15017.3		Q	B0JY34.1
S13	14290.4	-Met	A	B0JY40.1
S14	11836.6	-Met	A	B0JYC2.1
S15	10245.7	-Met	A	B0JPD9.1
S16	9561.1		I	B3DFB2.1
S17	8998.4	-Met	A	B0JHZ4.1
S18	8204.6	-Met	S	B0JVN4.1
S19	10128.8	-Met	G	B0JHZ9.1
S20	10420.1	-Met	A	B0JGK4.1
S21	7094.3	-Met	T	B0JIQ1.1
<b>Large Subunit Proteins</b>				
L06	19941.1	-Met	S	B0JHY9.1
L09	16662.3	-Met	A	B0JQ94.1
L13	16923.7		N	B0JY35.1
L14	13305.5		I	B0JHZ3.1
L15	15831.4		K	B0JHY6.1
L17	12964.2		R	B0JY37.1
L18	13250.1		K	B0JHY8.1
L19	13847.3		K	B0JNQ1.1
L21	14516.4	-Met	S	B0JI00.1
L22	13232.4	-Met	T	B0JHZ8.1
L23	11257.3	-Met	V	B0JI01.1
L24	12840.9	-Met	S	B0JHZ2.1
L28	8971.5	-Met	S	B0JLJ8.1
L29	8231.6	-Met	A	B0JHZ5.1
L31	8751.1	-Met	P	B0JY33.1
L32	6719.5	-Met	A	B3DFA5.1
L33	7270.4	-Met	A	B0JVN3.1
L34	5235.2	-Met	S	B0JN63.1
L35	7866.4	-Met	P	B0JSJ9.1
L36	4399.4		K	B3DFA8.1

doi:10.1371/journal.pone.0156275.t001

### Polygenetic analysis of *M. aeruginosa* strains

At present, the genomes of 12 *M. aeruginosa* strains have been completely or partially sequenced. As shown in [Table 2](#), the molecular weights of some of the ribosomal proteins identified varied among different strains because the amino acid sequences of these ribosomal proteins differed. By identifying the ribosomal proteins from *M. aeruginosa* NIES-843 as Type I, other types of ribosomal proteins could be identified and placed in numerical order. Several types of ribosomal proteins have been identified among these different strains.

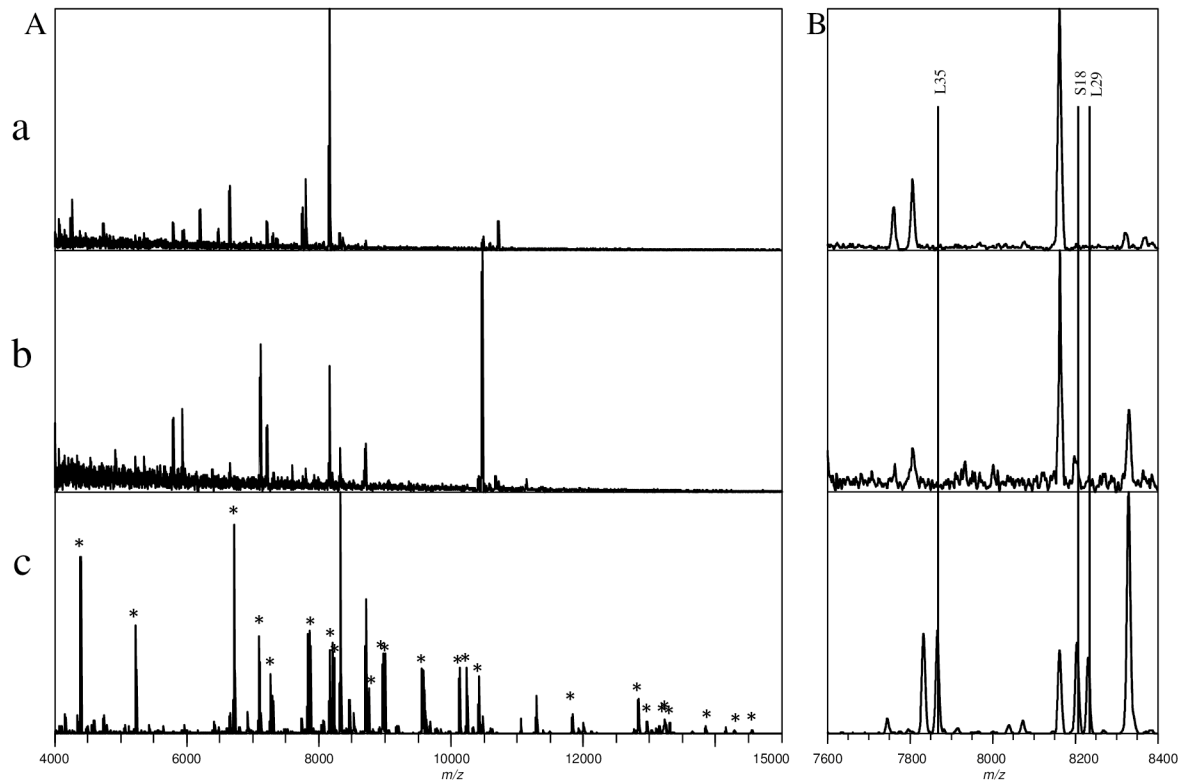
Thirteen of the 31 ribosomal proteins in the NIES-843 strain (L32, S21, L33, L35, S18, L29, L31, L28, S16, S19, S15, S20, S14)—those that were frequently detected in all 55 strains and appeared at high intensities between  $m/z$  4000–12000—were selected as biomarkers and classified by type, as summarized in [Table 2](#).



**Fig 1. Mass spectra of the *M. aeruginosa* (NIES-843) ribosomal proteins and their peak assignments.** Peaks that are identified as ribosomal proteins are marked with the protein's name.

doi:10.1371/journal.pone.0156275.g001

Fig 3 shows the expanded mass spectra ( $m/z$  10000–10550) of 4 *M. aeruginosa* strains (NIES-834, NIES-1070, NIES-109, and NIES-1216) as an example. These strains were selected from groups A, B, C, and D, respectively, as reported by Tanabe using MLST [18]. Ribosomal proteins S19, S15, and S20 are located in this mass spectrum range. Different versions of these three ribosomal proteins were identified in the mass spectra of the four example strains noted above. S19 exhibited the same mass (Type I) in all four strains. There were two forms of S15 among the four strains (Type I in NIES-843 and NIES-1216, and Type II in NIES-1070 and NIES-109). S20 exhibited more diversity among the four strains (Type I in NIES-843 and NIES-1070, Type IV in NIES-109, and Type V in NIES-1216). These results reveal that the four strains selected can be classified into different categories using the differences in their ribosomal proteins (Table 3).



**Fig 2. Comparison of the spectra of *M. aeruginosa* (NIES-843) samples prepared using different pre-treatment methods.** a: intact cell; b: cell lysate; c: concentrated ribosomal proteins. A: mass spectra ( $m/z$  4000–15000) of three pre-treatment methods. Peaks marked by \* are ribosomal proteins identified from concentrated ribosomal sample; B: Expanded view of the three pre-treatment methods' mass spectra ( $m/z$  7600–8400), in which ribosomal proteins L35, S18 and L29 were identified from the concentrated ribosomal sample.

doi:10.1371/journal.pone.0156275.g002

All 55 strains were classified following the above pattern using the different types identified for the 13 selected biomarkers, and a ribosomal protein-profiling table (S2 Table) was prepared. The ribosomal protein-profiling table was then processed using UPGMA cluster analysis, and a polygenetic tree was created, as shown in Fig 4.

Based on the differentiation of the 13 biomarkers by type, the 55 strains were categorized into 47 unique strain types, highlighting the high diversity of the *M. aeruginosa* strains. These strain types were divided into five major clades (I–V) using a similarity cut-off of 50%. Notably, Clades I and II were dominated by strains that possessed *mcyG*, although several exceptions were seen. Clades III, IV, and V included strains that did not possess *mcyG*. Thus, the toxic strains (which possessed *mcyG*, labeled in bold by strain name and indicated by a t in Fig 4) and non-toxic strains (which did not possess *mcyG*, indicated with an n in Fig 4) were divided into different groups based on their ribosomal protein types, as determined using MALDI-TOF MS. These results are very similar to the results of an MLST analysis previously used to distinguish toxic and non-toxic strains of cyanobacteria [18].

## Discussion

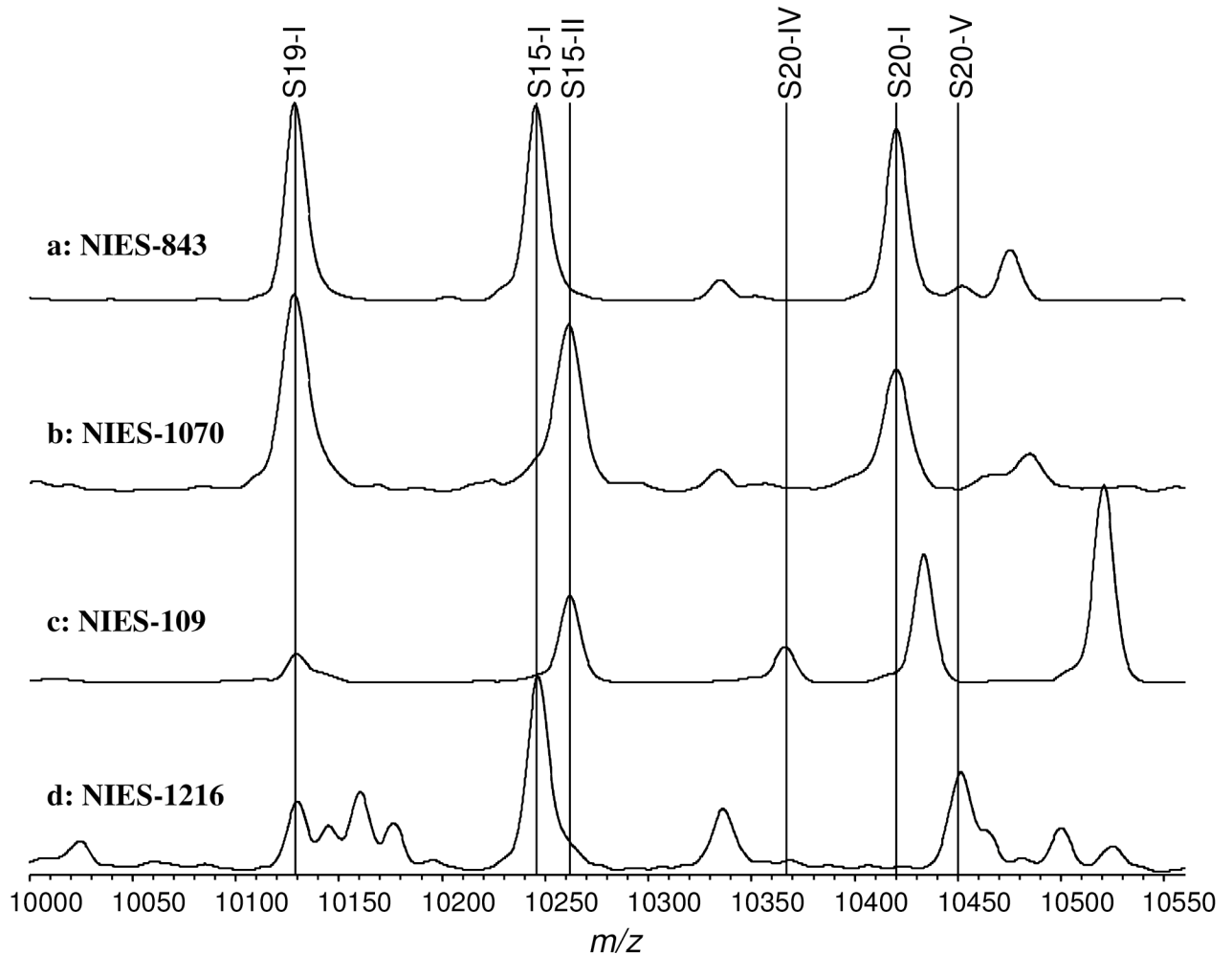
In this study, 31 of the 52 ribosomal proteins from *M. aeruginosa* NIES-843 were successfully identified using MALDI-TOF MS, indicating that by employing ribosomal proteins as biomarkers, cyanobacteria can be identified using MALDI-TOF MS.

**Table 2. Ribosomal protein list for 12 sequenced strains.**

Ribosomal proteins	MW (mass of [M+H])	PCC 7941	PCC 9432	PCC 9443	PCC 9701	PCC 9806	PCC 9808	THAIHU98	DIANCHI905	PCC 9717	PCC 9807	PCC 9809
L32	6719.5	I (6719.5)	-	-	II (6647.39)	I (6719.5)	-	-	-	-	-	-
S21	7094.3	I (7094.3)	I (7094.3)	I (7094.3)	I (7094.3)	I (7094.3)	I (7094.3)	-	-	-	-	-
L33	7270.4	I (7270.4)	I (7270.4)	I (7270.4)	I (7270.4)	I (7270.4)	-	I (7270.4)	-	-	-	-
L35	7866.4	II (7844.29)	III (7843.31)	III (7843.31)	I (7866.4)	I (7866.4)	III (7843.31)	-	-	-	I (7866.4)	-
S18	8204.6	I (8204.6)	I (8204.6)	I (8204.6)	I (8204.6)	I (8204.6)	I (8204.6)	-	I (8204.6)	-	-	-
L29	8231.6	II (8231.6)	II (8291.56)	III (8318.62)	IV (8305.58)	III (8318.62)	II (8291.56)	VII (8247.55)	II (8291.56)	V (8290.57)	III (8318.62)	VI (8261.57)
L31	8751.1	I (8751.1)	II (8737.07)	II (8737.07)	III (8765.13)	I (8751.1)	II (8737.07)	-	-	-	-	-
L28	8971.5	II (8971.5)	I (8971.5)	III (8957.46)	I (8971.5)	I (8971.5)	I (8971.5)	-	-	-	-	-
S16	9561.1	I (9561.1)	I (9561.1)	I (9561.1)	II (9531.04)	II (9531.04)	I (9561.1)	-	-	-	-	-
S19	10128.8	I (10128.8)	I (10128.8)	I (10128.8)	I (10128.8)	I (10128.8)	I (10128.8)	-	-	-	-	-
S15	10245.7	I (10245.7)	II (10261.73)	II (10261.73)	I (10245.7)	III (10275.76)	II (10261.73)	I (10245.7)	II (10261.73)	I (10245.7)	II (10261.73)	I (10245.7)
S20	10420.1	I (10420.1)	II (10436.06)	III (10539.14)	IV (10367.02)	V (10450.09)	VI (10422.03)	-	-	-	-	-
S14	11836.6	I (11836.6)	II (11806.54)	II (11806.54)	II (11806.54)	II (11806.54)	II (11806.54)	-	-	-	-	-

doi:10.1371/journal.pone.0156275.t002





**Fig 3. Expanded view of ribosomal proteins S19, S15, and S20 ( $m/z$  10000–10500) from four strains.** a: NIES 843; b: NIES 1070; c: NIES 109; d: NIES 1216. S19, S15, and S20 types are indicated by name.

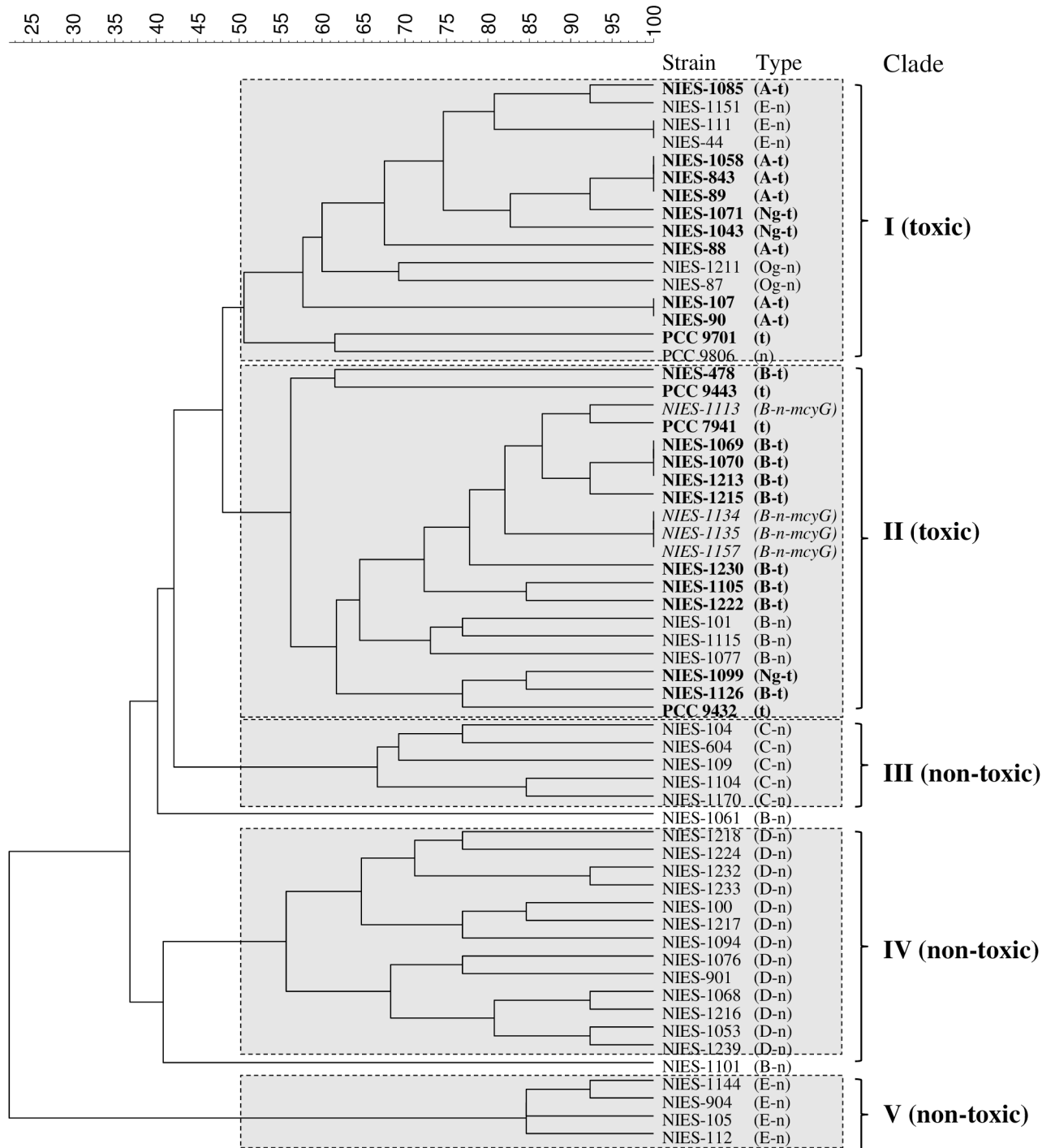
doi:10.1371/journal.pone.0156275.g003

It has been reported that ribosomal proteins can be detected directly from whole-cell samples of Gram-negative bacteria [35]. Although *M. aeruginosa* is a Gram-negative bacterium, it has a special cell wall and extracellular covering that includes a relatively thick peptidoglycan layer sandwiched between outer and inner membranes, as well as a coating of polysaccharide fibrils external to the outer membrane [36]. This complex cell wall structure may prevent the release of cellular contents and impede the effective ionization of ribosomal protein fractions.

**Table 3. Different S19, S15 and S20 ribosomal protein types for four strains.**

Strain	Ribosomal protein types		
	S19	S15	S20
NIES-843	1	1	1
NIES-1070	1	2	1
NIES-109	1	2	4
NIES-1216	1	2	5

doi:10.1371/journal.pone.0156275.t003



**Fig 4. Polygenetic tree of *M. aeruginosa* strains based on MALDI-TOF MS ribosomal protein profiling.** Strains are labeled and classified by their group and toxicity. The group names indicated in capital letters (A, B, C, D, E) were described by Tanabe (2007) using the MLST method. t: toxic; n: non-toxic; Ng-no group; Og: other group. In addition to the measured strains, 5 sequenced strains (PCC 7941, PCC9432, PCC 9443, PCC 9701, PCC 9806) were employed as references.

doi:10.1371/journal.pone.0156275.g004

The successful detection and identification of ribosomal proteins from the concentrated ribosome fraction demonstrates that a two-step pre-treatment procedure is an effective method to prepare *M. aeruginosa* samples for ribosomal protein identification. Because the process included only cell lysis and centrifugation, it can be considered a simple, rapid study method.

Thus far, identification methods related to *M. aeruginosa* could only be used to identify the genus and species, even when 16S rRNA gene analysis was employed. The present results suggest that combining high resolution MALDI-TOF MS and using ribosomal proteins as biomarkers provides an accurate method to distinguish multiple strains of *M. aeruginosa*. Further, these results prove that differences in *M. aeruginosa* can be distinguished at strain level.

In the primarily toxic Clade II (Fig 4), four strains (NIES-1113, NIES-1134, NIES-1135, and NIES-1157) that possessed *mycG* but did not produce MCs were identified. According to Rantala *et al.* [17], MC synthetase genes were originally present in all toxic strains of *Microcystis*, but some of these strains have lost the ability to produce MCs. Thus, the lack of toxicity in the four strains included in Clade II may stem from the fact that these strains lost their ability to produce MCs during evolution. The fact that these four strains fell into a toxic group and not a non-toxic group indicates the high discrimination ability of the MALDI-MS method and emphasizes the consistency between the results found in this study using ribosomal protein profiling and those found using *mycG* gene-level analysis. Furthermore, three of these strains formed a monophyletic sub-clade, and the last strain was located quite near this clade. This may indicate that *mycG* retained in these strains acts differently compared to other MC-producing genes.

Several non-toxic strains that did not possess *mycG* (NIES-1151, NIES-111, NIES-44, NIES-1211, NIES-87, NIES-101, NIES-1115, and NIES-1077) were also grouped in the otherwise toxic Clades (I and II). Rantala *et al.* [17] have suggested that the phylogenetic marker 16S rRNA gene and the microcystin synthetase gene are congruent; as a result, the fact that these strains do not possess the *mycG* gene but still fall into a toxic clade may imply a relative relationship to those strains that possess *mycG*.

Without exception, all three non-toxic clades (Clade III-V) exclusively contained strains that do not possess the *mycG* gene. This further demonstrates the high level of consistency between ribosomal protein profiling results and the MCs gene-typing process results. These results are significant because they demonstrate that toxic strains are not improperly identified when the method is applied to distinguish toxic and non-toxic strains.

The profiling of ribosomal proteins directly divided toxic and non-toxic *M. aeruginosa* strains into different clades. This suggests that measuring the ribosomal proteins of unknown strains or samples from natural waterbodies can be used to locate those strains within a toxic or non-toxic clade and provides a novel method to distinguish toxic *M. aeruginosa* strains within natural cyanobacterial samples. By continuing to analyze the mass spectrum database of *M. aeruginosa* strains, the accuracy of this method can be further improved.

Differentiating between toxic and non-toxic cyanobacteria strains is important to environmental management professionals because it can provide guidance on how to control these strains. Furthermore, the detection of toxic *M. aeruginosa* strains can help predict the release of MCs and allow for evaluations of the ecological risk of *M. aeruginosa* blooms in eutrophic waterbodies.

The polygenetic diversity of *M. aeruginosa* strains based on their ribosomal proteins is consistent with the MLST results presented by Tanabe [18] at the genetic level, which illustrates that ribosomal proteins are powerful markers in the study of polygenetic diversity. Theoretically, there are 52 ribosomal proteins in *M. aeruginosa* encoded by 52 genes. These proteins may provide additional genetic information beyond the seven housekeeping loci presented in Tanabe's study. The study of these proteins only requires a two-step pre-treatment and MS

measurement, which is much less time- and labor-intensive than previous strain analysis methods.

To our knowledge, the use of MALDI-TOF MS to analyze ribosomal proteins as biomarkers and distinguish toxic from non-toxic *M. aeruginosa* strains is a new concept and technique in the study of cyanobacteria. The simple pre-treatment and measurement process presented here enables the analysis of numerous samples in a short time period, such as in multi-point monitoring along or around a bloom's breakout site. Furthermore, because MALDI-TOF MS is an analytical chemistry method, the process can be automated and standardized. Thus, it has the potential to be applied in routine monitoring of cyanobacterial water blooms.

For a given species of cyanobacteria, the potential toxicity of a proliferating population varies greatly not only depending on the location of a bloom but also how long that bloom exists [37, 38]. Although no such data are available in most eutrophic waterbodies, the MALDI-TOF MS method presented here may allow for an analysis of spatiotemporal changes in the toxicity of cyanobacterial blooms, particularly given its ability to rapidly and accurately predict toxicity.

Future studies may address the parameters linking the characteristics of a cyanobacterial bloom, including local biological, physicochemical, and weather conditions, to that bloom's strain type. Understanding these conditions may provide insight into the mechanism of water bloom formation and subsequent control techniques. We plan to apply the results of this study to distinguish toxic strain types in natural samples and compare strain types from different eutrophic lakes, such as Kasumigaura Lake in Japan and Lake Taihu in China, in order to reveal the diversity of *M. aeruginosa* strains in different geographical regions.

## Supporting Information

**S1 Table. *M. aeruginosa* strain information.**

(PDF)

**S2 Table. Ribosomal protein peak typing of 55 strains.**

(PDF)

## Acknowledgments

The present study was supported by funding from the Jiangsu Environmental Monitoring Foundation, China (Project 1411) to Li-Wei Sun.

We would like to express our gratitude to Dr. Kosei Yumoto at MCC-NIES, Japan for providing cultures of *M. aeruginosa* strains. We thank Dr. Noriko Takamura for her collaboration in the present research.

## Author Contributions

Conceived and designed the experiments: LWS WJJ HS MK. Performed the experiments: LWS WJJ HS MK. Analyzed the data: LWS HS XWL. Contributed reagents/materials/analysis tools: LWS WJJ HS MK. Wrote the paper: LWS WJJ HS XWL.

## References

1. Francis G. Poisonous Australian Lake. *Nature*. 1878; 18: 11–12.
2. Ransom R, Soong FS, Fitzgerald J, Turczynowicz L, El Saadi O, Roder D, Maynard T, Falconer I. Health effects of toxic cyanobacteria (Blue-Green Algae). National Health and Medical Research Council, Canberra: Australian Government Publishing Service; 1994. p108.
3. Kuiper-Goodman T, Falconer I, Fitzgerald J. Toxic cyanobacteria in water: A guide to their public health consequences, monitoring, and management. Chorus I, Bartram J, editors. London: WHO E & FN Spon Publishers; 1999. P. 113–153.

4. World Health Organization. Guidelines for drinking water quality: Addendum to Volume 2. 2nd Ed. Geneva: Health criteria and other supporting information; 1998.
5. Lawton L. Determination of cyanobacteria in the laboratory. In: Chorus I, Bartram J, editors. Toxic cyanobacteria in water: A guide to their public health consequences, monitoring and management. London: WHO E & FN Spon Publishers; 1999. P. 347–367.
6. Strickland JDH, Parsons TR. Pigment analysis. A practical handbook of seawater analysis. Bulletin, Fisheries Research Board of Canada, Ottawa. 1968. p185–205.
7. Yentsch CS, Menzel DW. A method for the determination of phytoplankton chlorophyll and phaeophytin by fluorescence. Deep Sea Res Oceanogr Abstr. 1963; 10: 221–231.
8. Simis SGH, Peters SWM, Gons HJ. Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. Limnol Oceanogr. 2005; 50: 237–245.
9. Neilan BA, Jacobs D, Del DT, Blackall LL, Hawkins PR, Cox PT, et al. rRNA sequences and evolutionary relationships among toxic and nontoxic cyanobacteria of the genus *Microcystis*. Int J Syst Bacteriol. 1997; 47: 693–697. PMID: [9226902](#)
10. Tillett D, Parker DL, Neilan BA. Detection of toxigenicity by a probe for the microcystin synthetase A gene (*mcyA*) of the cyanobacterial genus *Microcystis*: Comparison of toxicities with 16S rRNA and phycocyanin operon (phycocyanin intergenic spacer) phylogenies. Appl Environ Microbiol. 2001; 67: 2810–2818. PMID: [11375198](#)
11. Neilan BA. Identification and phylogenetic analysis of toxigenic cyanobacteria by multiplex randomly amplified polymorphic DNA PCR. Appl Environ Microbiol. 1995; 61: 2286–2291. PMID: [16535049](#)
12. Lyra C, Suomalainen S, Gugger M, Vezie C, Sundman P, Paulin L, et al. Molecular characterization of planktic cyanobacteria of *Anabaena*, *Aphanizomenon*, *Microcystis* and *Planktothrix* genera. Int J Syst Evol Microbiol. 2001; 51: 513–526. PMID: [11321098](#)
13. Janse I, Kardinaal WE, Meima M, Fastner J, Visser PM, Zwart G. Toxic and nontoxic *Microcystis* colonies in natural populations can be differentiated on the basis of rRNA gene internal transcribed spacer diversity. Appl Environ Microbiol. 2004; 70: 3979–3987. PMID: [15240273](#)
14. Otsuka S, Suda S, Li R, Watanabe M, Oyaizu H, Matsumoto S, et al. Phylogenetic relationships between toxic and non-toxic strains of the genus *Microcystis* based on 16S to 23S internal transcribed spacer sequence. FEMS Microbiol Lett. 1999; 172: 15–21. PMID: [10079523](#)
15. Kurmayer R, Dittmann E, Fastner J, Chorus I. Diversity of microcystin genes within a population of the toxic cyanobacterium *Microcystis* spp. in Lake Wannsee (Berlin, Germany). Microb Ecol. 2002; 43: 107–118. PMID: [11984633](#)
16. Nishizawa T, Asayama M, Fujii K, Harada K, Shirai M. Genetic analysis of the peptide synthetase genes for a cyclic heptapeptide microcystin in *Microcystis* spp. J Biochem. 1999; 126: 520–529. PMID: [10467167](#)
17. Anne R, David PF, Michael H, Leo R, Jaana V, Thomas B, et al. Phylogenetic evidence for the early evolution of microcystin synthesis. PNAS. 2004; 101: 568–573. PMID: [14701903](#)
18. Yuuhiko T, Fumie K, Makoto MW. Multilocus sequence typing (MLST) reveals high genetic diversity and clonal population structure of the toxic cyanobacterium *Microcystis aeruginosa*. Microb. 2007; 153: 3695–3703.
19. Arnold RJ, Reilly JP. Observation of *Escherichia coli* ribosomal proteins and their posttranslational modifications by mass spectrometry. Anal Biochem. 1999; 269: 105–112. PMID: [10094780](#)
20. Claydon MA, Davey SN, Edwards-Jones V, Gordon DB. The rapid identification of intact microorganisms using mass spectrometry. Nat Biotechnol. 1996; 14: 1584–1586. PMID: [9634826](#)
21. Fenselau C, Demirev PA. Characterization of intact microorganisms by MALDI mass spectrometry. J Mass Spectrom Rev. 2001; 20: 157–171.
22. Friedrichs C, Rodloff A C, Chhatwal GS, Schellenberger W, Eschrich K. Rapid identification of *Viridans streptococci* by mass spectrometric discrimination. J Clin Microbiol. 2007; 45: 2392–2397. PMID: [17553974](#)
23. Mazurek GN, Weissman DN. Discrimination of intact mycobacteria at the strain level: a combined MALDI-TOF MS and biostatistical analysis. Proteomics. 2006; 6: 6416–6425. PMID: [17109381](#)
24. Keys CJ, Dare DJ, Sutton H, Wells I, Lunt M, McKenna T, et al. Compilation of a MALDI-TOF mass spectral database for the rapid screening and characterization of bacteria implicated in human infectious diseases. Infect Genet Evol. 2004; 4: 221–242. PMID: [15450202](#)
25. Krishnamurthy T, Ross PL. Rapid identification of bacteria by direct matrix-assisted laser desorption/ionization mass spectrometric analysis of whole cells. Rapid Commun Mass Spectrom. 1996; 10: 1992–1996. PMID: [9004531](#)

26. Lay JO. MALDI-TOF mass spectrometry of bacteria. *Mass Spectrom Rev.* 2001; 20: 172–194. PMID: [11835305](#)
27. Pineda FJ, Antoine MD, Demirev PA, Feldman AB, Jackman J, Longenecker M, et al. Microorganism identification by matrix-assisted laser/desorption ionization mass spectrometry and model-derived ribosomal protein biomarkers. *Anal Chem.* 2003; 75: 3817–3822. PMID: [14572048](#)
28. Rhyzov V, Fenselau C. Characterization of the protein subset desorbed by MALDI from whole bacterial cells. *Anal Chem.* 2001; 73: 746–750. PMID: [11248887](#)
29. Sun LW, Kanae T, Hiroaki S, Masaki T, Hiroaki T, Tomoyoshi S. Characterization of ribosomal proteins as biomarkers for matrix-assisted laser desorption/ionization mass spectral identification of *Lactobacillus plantarum*. *Rapid Commun Mass Spectrom.* 2006; 20: 3789–3798. PMID: [17120274](#)
30. Kanae T, Hiroaki S, Sun LW, Masaki T, Hiroaki T. A simple intact protein analysis by MALDI-TOF-MS for characterization of ribosomal proteins of two genome-sequenced lactic acid bacteria and verification of their amino acid sequences. *J Proteome Res.* 2007; 6: 3899–3907. PMID: [17854216](#)
31. Kanae T, Hiroaki S, Sun LW, Masaki T, Hiroaki T, Hiromichi Y, et al. Phylogenetic classification of pseudomonas putida strains by MALDI-TOF-MS using ribosomal subunit proteins as biomarkers. *Anal Chem.* 2007; 79: 8712–8719. PMID: [17939647](#)
32. Kasai F, Kawachi M, Erata M, Watanabe MM. NIES-Collection, List of strains, microalgae and protozoa. 7th ed. Tsukuba: National Institute for Environmental Studies; 2004.
33. Takakazu K, Nobuyoshi N, Shinobu O, Iwane S, Yuuhiko T, Masanori T, et al. Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. *DNA Res.* 2007; 14: 247–256. doi: [10.1093/dnares/dsm026](#) PMID: [18192279](#)
34. Moerschell RP, Hosokawa Y, Tsunasawa S, Sherman F. The specificities of yeast methionine aminopeptidase and acetylation of amino-terminal methionine in vivo-processing of altered iso-1-cytochromes-cis created by oligonucleotide transformation. *J Biol Chem.* 1990; 265: 19638–19643. PMID: [2174047](#)
35. Randy JA, James PR. Fingerprint matching of *E. coli* strains with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of whole cells using a modified correlation approach. *Rapid Commun Mass Spectrom.* 1998; 12: 630–636. PMID: [9621446](#)
36. Linda E G, James MG, Lee WW. ALGAE. 2nd ed. USA: PEARSON; 2009. P. 96–97
37. Kardinaal WEA, Visser PM. Dynamics of cyanobacterial toxins: sources of variability in microcystin concentrations, In: Huisman J, Matthijs HCP, Visser PM, editors. Harmful Cyanobacteria. Berlin: Springer; 2005. P. 41–43.
38. Enora B, Nicolas E, Cecile S, Marion S, Catherine Q, Jean-Francois H. Spatiotemporal changes in the genetic diversity of a bloom-forming *Microcystis aeruginosa* (cyanobacteria) population. *The ISME Journal.* 2009; 3:419–429 doi: [10.1038/ismej.2008.121](#) PMID: [19092863](#)