Water Research 203 (2021) 117543

Contents lists available at ScienceDirect

Water Research

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

Rapid and improved identification of drinking water bacteria using the Drinking Water Library, a dedicated MALDI-TOF MS database

Anna Pinar-Méndez [a,d,e,1,*], Sonia Fernández [b,1], David Baquero [b], Carles Vilaró [a], Belén Galofré [a], Susana González [b], Lidia Rodrigo-Torres [c], David R. Arahal [c], M. Carmen Macián [c], María A. Ruvira [c], Rosa Aznar [c], Laia Caudet-Segarra [d,e], Laura Sala-Comorera [d,e], Francisco Lucena [d,e], Anicet R. Blanch [d,e], Cristina Garcia-Aljaro [d,e,*]

[a] Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l'Aigua, General Batet 1-7, Barcelona 08028, Spain
[b] Cetaqua, Water technology center, Cornellà de Llobregat, Spain
[c] Departamento de Microbiología y Ecología and Colección Española de Cultivos Tipo (CECT), Universitat de València, Valencia, Spain
[d] Department of Genetics, Microbiology and Statistics, University of Barcelona, Avinguda Diagonal, 643, 08028 Barcelona, Spain
[e] The Water Research Institute, University of Barcelona, Barcelona, Spain

ARTICLE INFO

Keywords:
Bacterial identification
MALDI-TOF MS
Drinking water
Distribution network
Bottled water
Spectra database

ABSTRACT

According to the European Directives (UE) 2020/2184 and 2009/54/EC, which establishes the sanitary criteria for water intended for human consumption in Europe, water suitable for human consumption must be free of the bacterial indicators *Escherichia coli, Clostridium perfringens* and *Enterococcus* spp. Drinking water is also monitored for heterotrophic bacteria, which are not a human health risk, but can serve as an index of bacteriological water quality. Therefore, a rapid, accurate, and cost-effective method for the identification of these colonies would improve our understanding of the culturable bacteria of drinking water and facilitate the task of water management by treatment facilities. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) is potentially such a method, although most of the currently available mass spectral libraries have been developed in a clinical setting and have limited environmental applicability. In this work, a MALDI-TOF MS drinking water library (DWL) was defined and developed by targeting bacteria present in water intended for human consumption. This database, made up of 319 different bacterial strains, can contribute to the routine microbiological control of either treated drinking water or mineral bottled water carried out by water treatment and distribution operators, offering a faster identification rate compared to a clinical sample-based library. The DWL, made up of 96 bacterial genera, 44 of which are not represented in the MALDI-TOF MS bacterial Bruker Daltonics (BDAL) database, was found to significantly improve the identification of bacteria present in drinking water.

1. Introduction

Water quality for human consumption is regulated by the European Directives (UE) 2020/2184 and 2009/54/EC, which cover tap water and natural mineral water, respectively. In both types of water, an absence of fecal bacterial indicators is an indicator of safety, and heterotrophic bacteria are enumerated to manage water quality (Bartram et al., 2003), serving as an indicator of the efficacy of water treatment processes, mainly disinfection. Though not directly a risk for human health, heterotrophic communities can influence water quality, as they may include opportunistic pathogens and can form biofilms on surfaces. The complex ecosystems of drinking waters also contain viable but non-culturable bacteria, which represent a minor fraction of the overall drinking water communities and are still highly underexplored.

A wide range of techniques are available for bacterial identification: they can be culture-based, involving the use of chromogenic media or biochemical testing, or molecular, such as high-throughput sequencing of the 16S rRNA gene. Differing considerably in performance and

* **Corresponding authors at:** Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l'Aigua, General Batet 1-7, Barcelona 08028, Spain and Department of Genetics, Microbiology and Statistics, University of Barcelona, Avinguda Diagonal, 643, 08028 Barcelona, Spain.
E-mail addresses: apinar@aiguesdebarcelona.cat (A. Pinar-Méndez), crgarcia@ub.edu (C. Garcia-Aljaro).
[1] Both the authors contributed equally to this work.

https://doi.org/10.1016/j.watres.2021.117543
Received 26 May 2021; Received in revised form 23 July 2021; Accepted 7 August 2021
Available online 12 August 2021

requirements, culture-based methods can be time-consuming, whereas molecular methods are more expensive and need specific technical skills and expertise. Proteomics offers a well-balanced approach through the application of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), which analyses the constantly expressed high-abundance proteins (mainly ribosomal) of a microbial cell. This promising tool has equivalent identification power to Sanger 16S rRNA gene sequencing and provides faster results (Sala-Comorera et al., 2016; Sárvári et al., 2018; Timperio et al., 2017). Moreover, according to previous studies (Tan et al., 2012), the use of MALDI-TOF can produce annually net savings of 87.8% in reagent costs compared to traditional methods, or up to 57.1% savings when including technologist and maintenance. Even though initial investment in the equipment is high, it is offset after a short period, typically three years at a reasonable use. Identification can be performed by non-skilled personnel and the time to obtain results is shortened by one working day compared to biochemical methods (Tsuchida et al., 2020). Meanwhile, in genomic technologies, a high expertise in molecular techniques is needed and also the results take longer, it depends on sequencing type, within few days to two weeks. Despite the power of genomics, some discrepancies have been observed between MALDI-TOF analysis of cultivable bacteria and 16S rRNA high-throughput sequencing since they are targeting different populations: heterotrophic populations versus total bacteria (Comorera et al., 2020). In this case, although genomics is thought to be more powerful, a lower diversity index was observed in samples analyzed by 16S rRNA-sequencing compared to heterotrophic bacteria by MALDI-TOF MS, probably because the dominance of certain species hindered the identification of minor genera by high-throughput sequencing. Furthermore, other biases including primer bias and the lack of a standardized methodology are still a major shortcoming which difficult its use in routine analysis (Boers et al., 2019).

On the other hand, current water quality regulations still rely on culture-based methods to assess water safety. In addition to the mandatory monitoring of fecal indicator bacteria and heterotrophic plate counts, a more in-depth characterization of heterotrophic bacteria by MALDI-TOF MS could provide a better understanding of these bacterial communities and improve water management and distribution.

Studies have demonstrated the reliability, speed, and easy-to-use features of MALDI-TOF MS in different areas, including food (Angelakis et al., 2011; Pavlovic et al., 2013; Yu et al., 2019), groundwater (Santos et al., 2017), wastewater (Eddabra et al., 2012), mining (Avanzi et al., 2017) or even in space craft and associated surfaces (Seuylemezian et al., 2018). Recent research (Sala-Comorera et al., 2017) showed the suitability of MALDI-TOF MS for the routine monitoring of heterotrophic bacteria in a drinking water treatment plant, as well as in mineral water (Sala-Comorera et al., 2020). However, there is general agreement that the resolving power of MALDI-TOF MS is currently limited by the lack of environmental spectra in commercial databases and there is a need to create in-house libraries with the bacterial spectra of interest (De Carolis et al., 2014; Kopcakova et al., 2014; Rahi et al., 2016; Santos et al., 2016; Seuylemezian et al., 2018). Another drawback of this approach is that identification can be affected by differences in analytical variables, i.e., the extraction method, culture conditions, matrix and database (Rahi et al., 2016; Ruelle et al., 2004).

Among the different MALDI sample preparation methods (direct transfer, formic acid extended direct transfer and acid/acetonitrile extraction), the formic acid extended direct transfer has been described as the most cost-effective and time-saving (Ghosh et al., 2015; Theel et al., 2012). However, other authors have achieved better results when using extraction methods in fungal cells (Chalupová et al., 2014) and Gram-positive bacteria (Alatoom et al., 2011), as the complex and rigid cell walls need extra lysis to access their proteins.. In addition, manufacturer-recommended protocols may also vary depend on their ability to inactivate pathogenic organisms. Safety measures with respect to handling MALDI-TOF MS samples, requires at least work under a biosafety level 2 conditions (BSL-2) and extraction procedures reducing risk of exposure to potential pathogens. According to some studies, certain chemical or physical treatment used in sample preparation before processing on MALDI-TOF MS, contribute to biological inactivation of samples: e.g. trifluoracetic acid (TFA) inactivates potential pathogens especially bacterial endospores (Lasch et al., 2008; Drevinek et al., 2012), tube-based ethanol-formic acid-acetonitrile extraction followed by filtration is recommended for security-sensitive biological agents (Tracz et al., 2016) or 70% ethanol is suggested to be sufficient for non-spore forming bacteria inactivation (Cunningham and Patel, 2015). The choice of an appropriate extraction method is therefore crucial for good results and biological safety.

We here report, to our knowledge, the first MALDI-TOF MS database constructed specifically for the identification of bacteria present in water for human consumption, named the Drinking Water Library (DWL). It was created using spectra from a selection of related environmental reference strains from the Spanish Type Culture Collection (CECT), as well as drinking water isolates, properly identified by 16S rRNA gene sequencing. The spectra were generated by standardized procedures and compiled in a dedicated database. The DWL will assist the monitoring of culturable bacterial communities in drinking water by allowing a rapid and accurate identification of isolates. In addition, this study provides the basis to increase the DWL by further characterization of drinking water heterotrophic bacteria.

2. Materials and methods

2.1. Samples and sampling conditions

Two hundred and nine samples were taken of different types of water, including water from distribution networks (109), process water (68), spring water (2) and bottled natural mineral water (30). The samples were grouped into two categories based on their origin: treated (from distribution networks and process water) and non-treated (natural mineral water and spring water). Sampling and processing conditions were slightly different for each group, as described below.

For the treated water category, 109 chlorinated drinking water samples (0.2–1 ppm free chlorine) from distribution networks in Catalonia (North-East Spain), Andorra and South France were analyzed; 68 process water samples were collected at nine different stages of treatment in a drinking water treatment plant (DWTP) in Barcelona corresponding to river water, groundwater, sand filters, ozonation, granular activated carbon (GAC) filters, ultrafiltration, cartridge filters, reverse osmosis and chlorinated water. For all 177 samples, 1 liter of water was collected in polyethylene sterile bottles with sodium thiosulfate (24 mg/L) and transported to the laboratory at 4ºC for further analysis within 24 h.

For the non-treated water category, 12 different bottled water brands (natural mineral water and spring water) were selected and a total of 30 samples were analyzed. Natural mineral water samples consisted of water bottled in 1.5 L polyethylene terephthalate bottles acquired from different retailers and stored in the dark at room temperature ($22 \pm 2\,^{\circ}C$) until analysis within one month after bottling. In the case of spring water, two different sources were sampled in sterile 1 liter polyethylene bottles, kept at 4 °C and analyzed within 24 h.

2.2. Isolation of heterotrophic bacteria from treated and non-treated drinking water

Heterotrophic bacteria from drinking water samples were recovered by concentration through membrane filtration, hollow fiber membrane ultrafiltration, or mass inoculation.

2.2.1. Treated water

Out of 177 treated water samples, 112 were processed by membrane filtration. Volumes of 100 mL were filtered through 0.22 or 0.45 μm pore-sized mixed cellulose ester filters (Millipore, Germany). Filters

were incubated on Water Plate Count Agar ISO (WPCA) (Oxoid, Spain) at 22 ± 2 °C for 72 ± 3 h.

For the process water samples, as concentration methods for early-stage samples from the DWTP were not required, decimal serial dilutions between 0.01 mL and 1 mL were cultured by mass inoculation in WPCA at 22 ± 2 °C for 72 ± 3 h (ISO, 1999). Spread plating of the samples onto WPCA by a mass inoculation method was also done.

To recover a higher diversity in chlorinated matrices, larger volumes were sampled. Thus, 25 samples from distribution networks were filtered through the hollow fiber filter Rexeed™ 25-A (Asahi Kasei Medical Co, Japan), which is a haemodialysis ultrafilter typically used for clinical purposes but recently applied to concentrate water samples. The main advantage of this method is its capacity to concentrate bacteria, viruses and protozoa in large volumes of water (from 10 to 1000 L) (Hill et al., 2007; Gunnarsdottir et al., 2020). Briefly, samples ranging between 50 and 700 L were processed by this technique. After filtration and elution, 0.2 mL of the sample was inoculated onto WPCA by the spread plate technique and incubated at 22 ± 2 °C for 72 ± 3 h.

Those samples that showed no growth on WPCA were re-analyzed using R2A plates (Becton Dickinson, U.S.) at 22 °C for 120 ± 4 h. R2A agar is a low-nutrient medium reported to improve the recovery of stressed and chlorine-tolerant bacteria from drinking water systems (Reasoner and Geldreich, 1985). Colonies were picked up and subcultured on the corresponding culture media and they were further characterized by MALDI-TOF MS analysis.

2.2.2. Non-treated water

Samples from spring water and bottled natural mineral water were concentrated by membrane filtration. Different sample volumes (0.1 mL, 1 mL, 10 mL, 100 mL) were filtered through 0.22 μm pore size nitrocellulose membrane filters (Millipore, Germany) and incubated on WPCA (Oxoid, Spain) supplemented with 0.5 g/L 2,3,5-triphenyltetrazolium chloride (TTC) at 22 °C for 120 ± 3 h. Due to bacterial metabolism, colorless TTC is reduced to formazan, a red compound helping the visualization of the colonies.

Colonies were subcultured in WPCA for purification prior to MALDI-TOF MS analysis.

2.3. Identification of bacterial isolates by MALDI-TOF MS using the Bruker library

Bacterial isolates were obtained by subculturing one single well-isolated colony on new WPCA in order to achieve fresh pure cultures. Sample preparation for MALDI-TOF MS was carried out using the formic acid extended direct transfer method recommended by Bruker Daltonics. Thus, biological material (a single colony) from fresh agar cultures was smeared directly onto a spot on a MALDI target plate and then overlayed with 1 μL of 70% formic acid. After air-drying, 1 μL of matrix solution (saturated solution of a-cyano-4-hydroxycinnamic acid in a standard solvent (Sigma-Aldrich)) was added. The Bruker Bacterial Test Standard (Bruker Daltonics, Germany) containing an extract of *Escherichia coli* DH5 alpha peptide with a protein profile was included for each plate to calibrate the instrument and validate the run.

Mass spectra ranging from 2000 to 20,000 Da of each isolate were automatically acquired, using the Microflex LT MALDI-TOF MS device and MALDI BioTyper software, version 3.1 (Bruker Daltonics, Germany) with FlexControl software package (Bruker Daltonics, Germany). The obtained spectra were analyzed using MALDI Biotyper Real-Time Classification software package (Bruker Daltonics, Germany) with the bacterial Bruker Daltonics (BDAL) database MBT Compass Library DB-7311, and checked using the FlexAnalysis software (Bruker Daltonics, Germany). The Bruker identification results are classified with a score based on matches between the mass spectra of the analyzed sample and the mass spectra database. The identification categories were as recommended by Bruker: unreliable identification was in red (score ≤1.699); probable genus identification in yellow (score 1.700–1.999); secure genus and probable species identification in green (score 2.000–2.299); and highly probable species identification also in green (score ≥2.300–3.000).

2.4. Selection of isolates to be included in the Drinking Water Library

The DWL consisted of spectra obtained from strains isolated in different water matrices and the spectra of reference strains from the CECT. The selection procedure was as follows:

2.4.1. Selection of non-successfully identified isolates (red category) recovered from the different water matrices under study

Non-successfully identified isolates were grouped into similarity clusters, subcultured and reanalyzed using the acid/acetonitrile extraction method (Bruker Daltonics, Germany): ca. 10 mg biomass from agar cultures was first suspended in 300 mL water by careful mixing, and then the suspension was mixed with 900 mL ethanol. The biomass was collected by centrifugation and the pellet was re-suspended in 1–80 μL 70% formic acid. The suspension was mixed carefully with 1–80 μL acetonitrile. Immediately after centrifugation, the supernatant was removed and aliquots of 1 μL were placed on each spot of a stainless-steel target plate. After air-drying, 1 μL of matrix solution was added.

The isolates were analyzed, generating 9 spectra per isolate, which constitute the mini Main Spectrum Profile (mMSP), a useful approximation to simplify the clustering. Spectra were acquired, processed, and compared by MALDI BioTyper software, version 3.1 (Bruker Daltonics, Germany), and checked for quality using the FlexAnalysis software (Bruker Daltonics, Germany). All spectra were processed using the default program settings for smoothing (Savitzky-Golay), baseline subtraction (TopHat) and normalization. An averaged mass spectrum was created by eliminating those with the higher deviation. The parameters used were a mass range from 3000 to 10,000 Da and 500 ppm as a maximum error for the main spectrum profile (MSP). The mMSPs were incorporated into an *ad hoc* database for clustering purposes.

Clustering of non-identified bacterial isolates was carried out in two steps. First, a dendrogram was created by the standard MALDI Biotyper 3.1 MSP creation method (Bruker Daltonics, Germany), where distance values are relative and normalized to a maximum value of 1.000. Second, a comparison based on log score values where each mMSP was compared with the *ad hoc* mMSP database. On this basis, isolates representing each cluster were selected for identification by 16S rRNA gene sequencing analysis.

2.4.2. Selection of CECT reference bacterial strains

For the construction of the DWL database, reference environmental and water-related bacterial strains were selected from the CECT catalogue (Table S1). The selection was based on the source of isolation (environmental and water-related), growth temperature (20–30 °C) and growth culture medium (mostly WPCA and R2A) or other media such as GSP (Glutamate Starch Phenol-red agar, also known as *Pseudomonas Aeromonas* Selective Agar), Glycine Vancomycin Polymyxin Cycloheximide Agar, Xylose Lysine Deoxycholate agar and Brain Heart Infusion agar. Related species (same genus) to those selected were also included, even though they had not been isolated from the environment. All the selected reference strains were cultured according to the growth conditions described for each bacterial strain in the CECT catalogue and processed with the acid/acetonitrile extraction method recommended by Bruker Daltonics to obtain their MSP as described below in Section 2.6.

2.5. Identification of isolates selected as representative for DWL spectra

2.5.1. Sequencing of the 16S rRNA gene

Representative bacterial strains from each cluster were identified by partial 16S rRNA gene sequencing on a Genius thermocycler (Techne, Burlington, NJ, USA). The amplification mixture (50 μL) comprised 1 μL

(50 pmol/μL) each of universal primers amplifying a 1000-bp region of the 16S rRNA gene 616 V (forward): 5′-AGAGTTTGATYMTGGCTCAG-3′ and P699R (reverse): 5′- GGGTYKCGCTCGTTR-3′ (Integrated DNA Technologies), 0.25 μL (5 U/μL) of Taq DNA polymerase (Takara, Clontech Laboratories, Inc.), 5 μL of 10X reaction buffer, 4 μL of dNTP mixture (10 mM), 33.75 μL of sterile filtered water (Milli-Q purification system, Millipore, Billerica, MA, USA), and 5 μL of DNA template (40 ng/μL). The DNA templates were amplified by initial denaturation at 94ºC for 5 min, followed by 35 cycles of denaturation at 94ºC for 30 s, annealing at 56ºC for 30 s, extension at 72ºC for 45 s, and a final extension at 72ºC for 10 min. Negative controls, devoid of DNA, were simultaneously included in the amplification process.

PCR amplicons of the partial 16S rRNA gene were verified by visualization in 1% agarose gel electrophoresis. Amplicons were purified and sequenced by the Macrogen Company Inc. Madrid, Spain. Subsequent sequencing reactions were done on an Abi Prism 3700 automated sequencer using the Big Dye Terminator v3.1 cycle sequencing kit. Sequencing primers were the same ones as used in the amplification reaction but diluted five-fold (10 pmol/μL). The taxonomic classification of bacterial isolates was performed using EzTaxon (Kim et al., 2012; Yoon et al., 2017) and BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul et al., 1997). Identification criteria for species delimitation varied depending on the phyla, ranging from 98.2% for strains of the phylum *Bacteroidetes* to 99.0% for *Actinobacteria* and those sequences with less than the respective cut-off values were considered potential new species, as in Meier Kolthoff et al. (2013). All strains are currently deposited at the public catalogue of the CECT and their identification results are shown in Table S2.

2.5.2. GC-FAME analysis

Complementary to the 16S rRNA gene sequencing, GC-FAME (Gas Chromatographic Fatty Acid Methyl Esters) analysis of bacterial strains was carried out as a method of strain authentication. Fatty acid methyl esters were extracted from biomass grown for 48 h on WPCA (Oxoid, Spain) at 26ºC and prepared according to standard protocols as described for the MIDI Microbial Identification System (Sasser, 1990) at the CECT. Cellular fatty acid content was analyzed by GC with an Agilent 6850 chromatographic unit, with the MIDI Microbial Identification System using the RTSBA6 method (MIDI Inc, Newark, US)) and identified using the Microbial Identification Sherlock software package. Only for those strains that did not show growth under these conditions was the analysis performed at their optimal growth conditions.

2.6. Generation of reference MSP for inclusion in the DWL

Selected isolates were inoculated in WPCA at 22 ± 2 °C and fresh cultures were used for MSP creation. To generate MSPs for inclusion in the DWL, a total of eight independent replicates of each isolate were processed, obtaining a collection of 24 spectra, three per replicate. Sample preparation for MALDI-TOF MS was carried out using the acid/acetonitrile extraction method recommended by Bruker Daltonics, as described before in section 4.b. The spectra measurements were recorded using the UltrafleXtreme (Bruker Daltonics, Germany) at VISAVET (UCM, Madrid, Spain). The instruments are equipped with a nitrogen laser. All spectra were recorded in linear positive mode with an acceleration voltage of 20 kV. Spectra were acquired by accumulating 250 laser shots across a spot. A mass range of 2000–20,000 *m/z* was used for analysis. The mass spectrometer was externally calibrated with the Bruker Bacterial Test Standard (Bruker Daltonics, Germany). FlexAnalysis software (Bruker Daltonics, Germany) was used for baseline subtraction, smoothing of spectra and mass labelling of peaks. MSPs in the DWL were represented by 18–24 good quality spectra.

2.7. Validation of the DWL

To evaluate the DWL performance, five internal and four external validation exercises were carried out by different technicians and different MALDI-TOF spectrometers. Validation exercises consisted of blind studies for the identification of bacterial isolates whose MSP had been included in the DWL. For internal validation, selected strains were analyzed to evaluate the identification results based on different protein extraction methods: direct transfer, formic acid extended direct transfer and acid/acetonitrile extraction. Corresponding spectra were obtained with the online acquisition method Biotyper Real Time Classification and the Microflex mass spectrometer (Bruker Daltonics, Germany). For identification, DWL and BDAL libraries were used. In addition, six other research centers in Spain performed the external validation. Samples consisted of 1.5 mL microcentrifuge tubes of bacterial pellet resuspended in ethanol:water 1:3 and stored at −20ºC until analysis. In this trial, 30 strains were processed in parallel for MALDI-TOF MS analysis using an extraction method and different mass spectrometer devices: Ultraflex III (Bruker Daltonics, Germany), UltrafleXtreme (Bruker Daltonics, Germany) and Microflex (Bruker Daltonics, Germany). All spectra were obtained with the online acquisition method Biotyper Real-Time Classification (Bruker Daltonics, Germany). Matrix blaster function was previously performed by firing eight initial laser shots at 25% intensity to get rid of the very first layer on sample. Then, mass spectra were automatically acquired in steps of 50 shots for a total of 250 shots accumulated, with initial laser power at 15% and maximal at 25%. The results based on log score value and identification were compared with the DWL and BDAL libraries.

3. Results and discussion

3.1. Identification of isolates using the Bruker Daltonics (BDAL) database

A total of 3809 bacterial isolates were analyzed by MALDI-TOF MS using the BDAL database: 1085 from distribution networks, 442 from process water, 356 from spring water and 1926 from bottled natural mineral water.

Regarding the treated water samples, 288 isolates from distribution networks were identified up to species level (26.5%), 266 at the genus level (24.5%) and 531 were not reliably identified (49%). In process water samples, 143 out of 442 isolates were identified at the species level (32.4%), 115 at the genus level (26%) and 184 were unidentified (41.6%) (**Fig. 1A**).

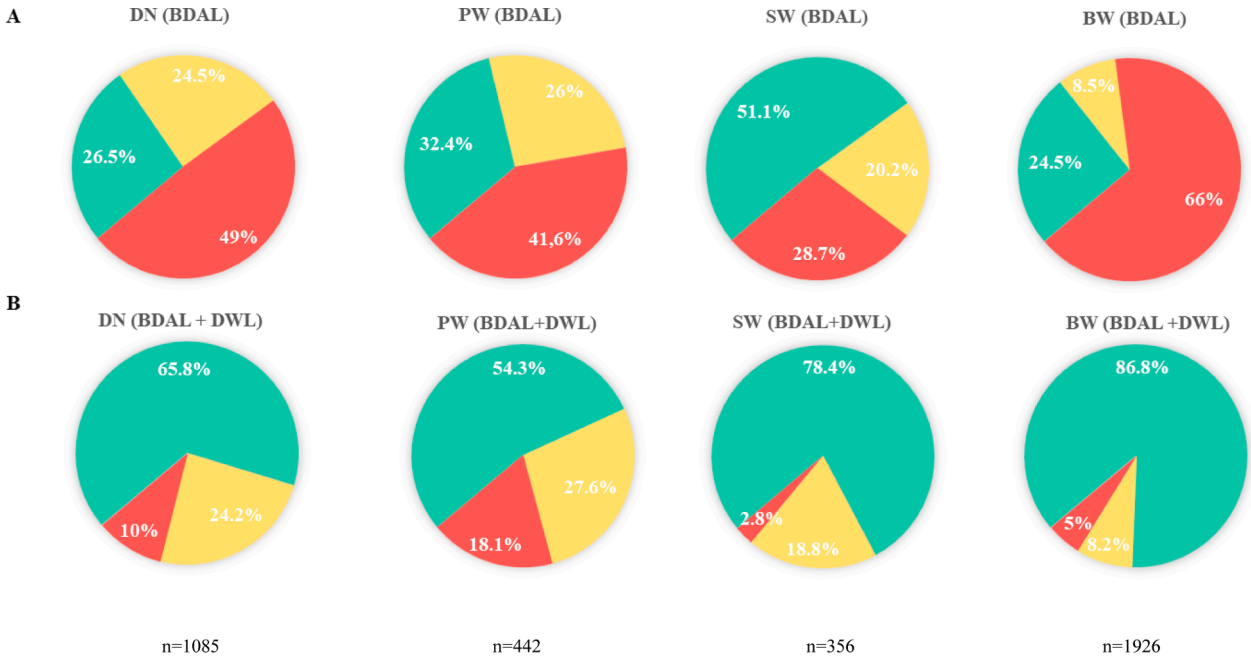
In the non-treated water samples, 182 out of 356 isolates (51.1%) from spring water were identified at the species level, 72 (20.2%) at the genus level and 102 (28.7%) remained unidentified. In bottled natural mineral water, 490 out of 1926 isolates (24.5%) were identified at the species level, 165 (8.5%) at the genus level, whereas 1271 isolates (66%) were unidentified, representing the highest number of unknown strains among the different samples (**Fig. 1A**).

In summary, 1721 isolates (45%) were successfully identified with the BDAL database at the species and/or genus level, whereas 2088 isolates remained unidentified (55%).

3.2. Identification of isolates not matched to the BDAL database

A total of 134 cluster-representative isolates were selected for identification by partial 16S rRNA gene sequencing. However, after preservation, only 120 grew satisfactorily and could be sequenced and further characterized by FAME profiles (data not shown), which established affiliation to 53 genera (**Table S2**). The isolates identified by 16S rRNA gene sequencing corresponded to distribution networks (45), process water (28), bottled natural mineral water (35) and spring water (12) (**Fig. 2B**). Therefore, the water from distribution networks accounted for the highest number of species not represented in the BDAL database.

Fourteen out of the 53 identified genera were absent in the BDAL database. In addition, some isolates (67) belonged to species not included in the BDAL database or to yet-to-be described. Identification of non-matched isolates by 16S rRNA sequencing allowed the BDAL



	n=1085			n=442			n=356			n=1926		
	species	genus	unrealiable	species	genus	unrealiable	species	genus	unrealiable	species	genus	unrealiable
BDAL	288	266	531	143	115	184	182	72	102	490	165	1271
BDAL + DWL	714	263	108	240	122	80	279	67	10	1671	157	98

Fig. 1. Percentage of isolates identified using the original Bruker Daltonics (BDAL) database (A) and the BDAL database expanded with the Drinking Water Library (DWL) (B) from different sources of water: distribution networks (DN), process water (PW), spring water (SW) and bottled natural mineral water (BW). Number of isolates (n) per source and according to identification results are shown in the table below the graph. When using the extended database, better identification results were obtained at the level of species (green) and genus (yellow), and unreliable identification (red) was reduced.

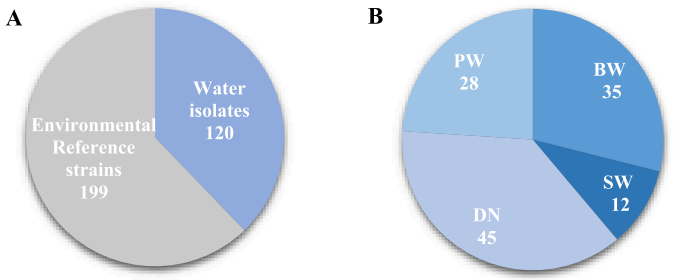


Fig. 2. Contents of the Drinking Water Library (DWL) and origin of the water isolates. (A) Composition of the 319 MSPs of the DWL according to the number of bacterial strains included in the database from environmental reference strains and water isolates. (B) Number of bacterial strains included in the new database according to origin: process water (PW), distribution network (DN), bottled mineral water (BW) and spring water (SW).

database to be expanded after the incorporation of their MSPs to the DWL (Table 1). A high variability in origin was observed among these isolates; some of the genera were recovered from multiple sources and others from only one. These results might be affected by a bias in the colony selection, as the main aim of the study was to recover the highest bacterial diversity from the four water sources. Nineteen genera were recovered from both treated (distribution network and process water) and non-treated water (spring water and bottled mineral water): *Acidovorax, Bacillus, Bosea, Brevundimonas, Caulobacter, Chryseobacterium, Deinococcus, Dyadobacter, Flavobacterium, Hydrogenophaga, Methylobacterium, Microbacterium, Mycobacterium, Paenibacillus, Pedobacter, Pseudomonas, Sphingobium, Sphingomonas* and *Sphingopyxis*. Nine genera were isolated only from distribution networks: *Acinetobacter, Blastomonas, Domibacillus, Ensifer, Micrococcus, Novosphingobium, Porphyrobacter, Pseudoxanthomonas* and *Shinella*. Eight genera were recovered only from process water: *Cloacibacterium, Ideonella, Massilia, Nocardioides, Novispirillum, Rheinheimera, Rhodoferax* and *Roseomonas*.

Similarly, for non-treated water samples, a total of six genera were exclusively recovered in spring water: *Aeromicrobium, Herminiimonas, Nocardia, Psychrobacillus, Williamsia* and *Xanthomonas*. And finally, regarding bottled natural mineral water samples, 11 different genera were identified as *Aquabacterium, Bradyrhizobium, Janibacter, Limnobacter, Methylibium, Pararhizobium, Phyllobacterium, Polaromonas, Psychrobacter, Rhizobium* and *Variovorax*.

3.3. The Drinking Water Library

The DWL constructed in this study (Fig. 2) includes 319 MSPs generated for isolates unmatched to the BDAL database and identified by 16S rRNA gene sequencing (120), representing the different clusters in which the isolates were grouped (**Table S2**). Also included are spectra from reference strains (199) selected for their environmental or water-related origins (**Table S1**).

The MALDI-TOF MS identification capacity increased when using the DWL with the newly incorporated MSPs, as half the genera in the library were absent from the BDAL database. At the beginning of the study in 2016, the BDAL database included 428 bacterial genera, which allowed the identification of 35% of drinking water isolates. The MSPs included in the DWL correspond to species belonging to 96 bacterial genera (Table 1), 44 of which were absent from the BDAL database and one was potentially a new genus. For the remaining 52 genera, 74 MSPs corresponding to one or more species were included, reaching a total of 319

Table 1
Contents of libraries BDAL 7311, DWL, and BDAL plus DWL according to the number of species per genera, and the origin of the main spectrum profiles (MSPs) that configured the new in-house DWL database (BW: Bottled mineral water, SW: Spring water, DN: Distribution network, PW: Process water and Ref: reference strains from the CECT catalogue). In brackets, number of new MSPs added. Genera not included in the BDAL database are shown in bold.

Genera included in the database	BDAL		DWL		BDAL+ DWL	Origin of MSP (DWL)				
	No. MSP	No. species	No. MSP	No. species	No. MSP	**BW**	**SW**	**DN**	**PW**	**Ref**
Acidovorax	9	7	+ (2)	1	11	X		X		
Acinetobacter	120	22	+ (3)	2	123			X		X
Aeromicrobium	1	1	+ (1)	1	2		X			
Aeromonas	46	18	+ (38)	9	84					X
Alcanivorax	–	–	+ (1)	1	1					X
Ampullimonas	–	–	+ (2)	1	2					X
Ancylobacter	–	–	+ (1)	1	1					X
Aquabacterium	–	–	+ (3)	3	3	X				
Arcicella	–	–	+ (1)	1	1					X
Arcobacter	15	6	+ (2)	2	17					X
Azoarcus	3	3	+ (1)	1	4					X
Azonexus	–	–	+ (1)	1	1					X
Bacillus	145	102	+ (4)	1	149			X	X	
Belliella	–	–	+ (1)	1	1					X
Blastomonas	2	2	+ (5)	4	7			X		X
Bosea	–	–	+ (5)	5	5	X		X	X	
Bradyrhizobium	2	2	+ (1)	1	3	X				
Brevundimonas	15	8	+ (9)	5	24	X		X		X
Caulobacter	2	2	+ (5)	4	7	X		X		X
Chitinimonas	–	–	+ (3)	2	3					X
Chromobacterium	–	–	+ (1)	1	1					X
Chryseobacterium	29	15	+ (22)	16	51	X		X	X	X
Cloacibacterium	–	–	+ (1)	1	1				X	
Cronobacter	9	1	+ (1)	1	10					X
Deinococcus	4	2	+ (2)	2	6			X	X	
Domibacillus	–	–	+ (1)	1	1			X		
Duganella	–	–	+ (1)	1	1					X
Dyadobacter	–	–	+ (2)	2	2			X	X	
Emticicia	–	–	+ (2)	2	2					X
Ensifer	–	–	+ (1)	1	1			X		
Ferruginibacter	–	–	+ (1)	1	1					X
Flavobacterium	15	14	+ (18)	12	33		X		X	X
Flectobacillus	–	–	+ (2)	2	2					X
Heliimonas	–	–	+ (2)	1	2					X
Herbaspirillum	17	11	+ (1)	1	18					X
Herminiimonas	2	2	+ (2)	2	4		X			
Hydrogenophaga	2	2	+ (2)	2	4	X			X	
Hymenobacter	–	–	+ (2)	2	2					X
Ideonella	–	–	+ (1)	1	1				X	
Iodobacter	1	1	+ (2)	1	3					X
Janibacter	2	2	+ (1)	1	3	X				
Janthinobacterium	1	1	+ (1)	1	2					X
Kinneretia	–	–	+ (1)	1	1					X
Klebsiella	31	3	+ (1)	1	32					X
Lacihabitans	–	–	+ (1)	1	1					X
Lactobacillus	249	97	+ (1)	1	250					X
Limnobacter	–	–	+ (1)	1	1	X				
Massilia	4	2	+ (4)	4	8				X	X
Methylibium	–	–	+ (1)	1	1	X				
Methylobacterium	22	11	+ (11)	8	33	X		X		X
Methylosinus	–	–	+ (2)	2	2					X
Microbacterium	55	38	+ (2)	2	57	X			X	
Micrococcus	21	4	+ (1)	1	22			X		
Mucilaginibacter	–	–	+ (10)	10	10					X
Mycobacterium	74	39	+ (7)	6	81		X	X		X
Nevskia	–	–	+ (2)	1	2					X
Nocardia	117	47	+ (1)	1	118		X			
Nocardioides	2	2	+ (1)	1	3				X	
Novispirillum	–	–	+ (1)	1	1				X	
Novosphigobium	18	12	+ (5)	4	23			X		X
Paenibacillus	137	69	+ (6)	5	143	X	X	X		X
Pararhizobium	1	1	+ (1)	1	2	X				
Parasediminibacterium	–	–	+ (1)	1	1					X
Pedobacter	2	2	+ (10)	10	12	X			X	X
Pelomonas	2	2	+ (1)	1	3					X
Pheaeospirillum	–	–	+ (1)	1	1					X
Phyllobacterium	–	–	+ (1)	1	1	X				
Polaromonas	–	–	+ (3)	2	3	X				
Porphyrobacter	–	–	+ (1)	1	1			X		
Propionivibrio	–	–	+ (1)	1	1					X
Pseudomonas	174	93	+ (10)	5	184		X	X		X

(*continued on next page*)

Table 1 (*continued*)

Genera included in the database	BDAL		DWL		BDAL+ DWL	Origin of MSP (DWL)				
	No. MSP	No. species	No. MSP	No. species	No. MSP	**BW**	**SW**	**DN**	**PW**	**Ref**
Pseudoxanthomonas	5	3	+ (2)	2	7			X		
Psychrobacillus	3	3	+ (1)	1	4		X			
Psychrobacter	3	2	+ (8)	5	11	X				X
Rheinheimera	1	1	+ (4)	4	5				X	X
Rhizobium	16	3	+ (1)	3	17	X				X
Rhodococcus	96	27	+ (5)	3	101			X		X
Rhodoferax	–	–	+ (1)	1	1				X	
Roseateles	–	–	+ (2)	2	2					X
Roseomonas	4	1	+ (2)	1	6				X	X
Runella	–	–	+ (2)	2	2					X
Sediminibacterium	–	–	+ (1)	1	1					X
Shinella	–	–	+ (1)	1	1			X		
Sphingobacterium	16	6	+ (1)	1	17					X
Sphingobium	15	11	+ (6)	6	21			X	X	
Sphingomonas	62	32	+ (10)	10	75	X		X		X
Sphingopyxis	5	5	+ (8)	8	15	X	X	X		X
Sphingorhabdus	–	–	+ (1)	1	1					X
Tabrizicola	–	–	+ (1)	1	1					X
Taeseokella	–	–	+ (1)	1	1					X
Undibacterium	–	–	+ (1)	1	1					X
Variovorax	7	1	+ (2)	2	9	X				
Vibrio	92	54	+ (2)	2	95					X
Williamsia	–	–	+ (1)	1	1		X			
Xanthomonas	36	17	+ (1)	1	37		X			
Yersinia	71	12	+ (1)	1	72					X
TOTAL	1783	824	319	240	2102	23	11	25	18	62

MSPs. The predominant genera were *Mucilaginibacter* with 10 MSPs, followed by *Bosea* with 5 MSPs and *Aquabacterium, Chitinimonas* and *Polaromonas* with 3 MSPs each. The newly added genera are shown in Table 1. Of the 44 new genera included in the DWL, 30 correspond to reference strains and 14 to isolates from treated water and non-treated water. At the species level, the DWL database provides 189 species not represented in the BDAL database, 67 of which are potentially new taxa.

Independently of their origin, whether environmental or reference strains, the new MSPs expanded the in-house database and improved the capacity of MALDI-TOF MS to identify drinking water isolates. Distribution of the DWL MSPs per genera and origin is shown in Table 1. Briefly, the genera included in the DWL were represented by variable numbers of MSPs. The most highly represented genus was *Aeromonas*, with 38 MSP entries (11.9%), followed by *Chryseobacterium* with 22 MSPs (6.9%), *Flavobacterium* with 18 MSPs (5.6%), *Sphingomonas* with 13 MSPs (4.1%) and *Methylobacterium* with 11 MSPs (3.4%). The genera *Pedobacter, Pseudomonas, Sphingopyxis* and *Mucilaginibacter* contributed 10 MSPs (3.1% each), whereas *Brevundimonas* provided 9 MSPs (2.8%), *Psychrobacter* 8 MSPs (2.5%), *Mycobacterium* 7 MSPs (2.2%) and *Sphingobium* and *Paenibacillus* 6 MSPs (1.9% each). The genera *Blastomonas, Rhodococcus, Bosea, Novosphigobium* and *Caulobacter* provided 5 MSPs (1.6% each) and the genera *Bacillus, Rheinheimera* and *Massilia* 4 MSPs (1.3% each).

3.4. Validation of the DWL database

The newly developed DWL was validated by (a) internal assays with reference CECT strains and drinking water isolates; (b) external blind trials; (c) re-identification of the drinking water isolates recovered in this project, and (d) identification of new isolates from routine laboratory analysis.

3.4.1. Internal assays

An internal assay was carried out with 9 CECT reference strains (CECT 317 *Brevundimonas diminuta,* CECT 7302 *Chryseobacterium aquaticum*, CECT 7791 *Flavobacterium tructae*, CECT 5998 *Methylobacterium aquaticum*, CECT 7550 *Mucilaginibacter myungsuensis*, CECT 7273 *Mycobacterium llatzerense*, CECT 153 *Paenibacillus polymyxa*, CECT 7114 *Pedobacter aquatilis* and CECT 8016 *Sphingopyxis italica*). The selection was based on the bacterial diversity found in drinking water and protein extraction difficulty in order to challenge the identification capacity of the new database. The strains were analyzed by MALDI-TOF MS, as described in Material and Methods (direct transfer, formic acid extended direct transfer and acid/acetonitrile extraction methods). Six out of the nine strains were successfully identified at the species level with scores of 2.04–2.55; two strains were identified at the genus level (1.73–1.94) and one was not identified (1.48–1.56) (**Table S3**). In general, differences between extraction methods were not observed, the score values being very close. Strain CECT 7273 *Mycobacterium llatzerenze* was only identified to the genus level by the extraction method (score 1.73–1.80). As described in other studies, due to their complex cell walls, *Mycobacterium* spp. require a specific extraction method for a better protein recovery, which may affect MALDI-TOF results (Alcolea-Medina et al., 2019). Strain CECT 153 *Paenibacillus polymyxa* was not identified (scores 1.49–1.56) when using the DWL but genus-level identification was achieved with the BDAL database (scores 1.84–1.98). It is well known that sporulation can affect identification by MALDI-TOF analysis, as spores differ in protein content compared to vegetative cells (Lasch et al., 2009).

3.4.2. External blind trials

For external validation, four rounds of proficiency testing of the DWL were carried out, each round involving at least three Spanish research centers and 29 strains (**Table S4**), which represents about 10% of the DWL content. One batch of cell culture was prepared per strain and distributed among the centers for the parallel analysis by MALDI TOF MS using the acid/acetonitrile extraction method and the mass spectrometer device available in the center (in all, three different instruments were used). As a result, 81% of measurements resulted in identification at probable species level (score >2.000) and 13% at probable genus level (score 1.700–1.999), 6% not leading to identification (score <1.699). In general, spectra were consistent, although occasional inter-laboratory and even intra-laboratory variability was observed. Differences in log score values between research centers can be explained either because of the cell composition (sporulated cultures, cell wall complexity), the elapsed time before sample processing, the conditions of sample preservation, or sample handling.

Likewise, inter-laboratory differences in analytical sensitivity were

registered for certain taxonomic groups that are nearly identical by 16S rRNA gene sequencing (Janda and Abbott, 2007). This was the case of strain *Flavobacterium* sp. CECT 9288, for which the highest log score value matched a closely related phylogenetic species of the genus *Flavobacterium*. As a result, two out of six centers only achieved identification at the genus level.

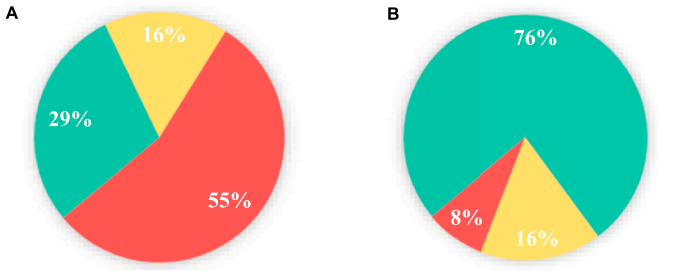
Slightly different results were observed depending on the spectrometer device used. For some strains, profiles clustered separately in Microflex vs UltrafleXtreme/Ultraflex II. Moreover, the quality of spectra also directly affected the results. The higher laser intensity of automatic compared to manual acquisition may produce background noise that impairs sensitivity, resulting in false or unreliable classification. For instance, one of the six laboratories in the trial acquired three spectra for strain CECT 9470 *Phyllobacterium myrsinacearum*, yet identification at species level was achieved with only one spectrum.

3.4.3. *Re-identification of the drinking water isolates recovered in this study*

The extended database (BDAL plus DWL) allowed offline re-analysis of isolates unidentified with previously acquired spectra. The performances of both databases are shown in Fig. 3. The BDAL database could identify 35% of isolates to levels of probable genus (score 1.700–1.999), secure genus and probable species (score 2.000–2.299), or highly probable species identification (score ≥2.300–3.000), which increased up to 95% when the DWL was used as well. The BDAL library alone failed to identify 2088 (54.8%) of the isolates, whereas 618 (16.2%) were identified at the genus level and 1103 (29%) at the species level. In contrast, when the same spectra were searched against the BDAL database plus the DWL, the identification performance improved by 47%, 2904 of the isolates being identified at the species level (76.2%) and 296 remaining unidentified (7.8%). The unidentified fraction corresponded to single representatives (non-clustered isolates) not included in the DWL as they were detected only once throughout the study.

In general, the addition of new MSP entries in the DWL increased the coverage of genus and species diversity, thereby improving the discrimination capacity of the original database. As a result (see Fig. 1**B**), 714 isolates from the distribution networks were successfully identified at the species level (65.8%), 263 at the genus level (24.2%) and 108 isolates remained unknown (10%). The results for process water samples were similar: from a total of 1085 isolates, 240 were identified at species level (54.3%), 122 at genus level (27.6%) and 80 remained unknown (18.1%). In the two groups, the percentage of unknown samples decreased by 38.9% and 23.5%, respectively, when using DWL. Surprisingly, this reduction was significantly greater for the bottled natural mineral water samples. When the 1926 spectra were searched against the extended database, 1671 were identified at the species level (86.8%) and another 157 (8.2%) at the genus level; only 98 isolates remained unidentified, which represented an improvement of 60.9% over the BDAL database (Fig. 1**A**). The extended database also gave better results in spring water samples: a total of 97.2% (346 out of 356) of isolates were successfully identified, 78.4% at species level and 18.8% at genus level, with only 10 isolates unidentified.

As shown, the DWL contributed new MPSs on environmental strains lacking in the original database and constitutes a good tool for the routine testing of drinking water samples. Although databases are frequently updated by the providers, results of the present study highlighted, in accordance with other authors (De Carolis et al., 2014; Kopcakova et al., 2014; Rahi et al., 2016; Seuylemezian et al., 2018) that there is a need for continuously updated in-house databases to increase taxonomic resolution, especially for microbial ecology studies.



n=3809		
species	genus	unrealiable
1103	618	2088

n=3809		
species	genus	unrealiable
2904	609	296

Fig. 3. The percentage of isolates identified using the original Bruker Daltonics (BDAL) database (A) compared with the BDAL database plus the Drinking Water Library (B). In green: highly probable species (score ≥2.300–3.000), secure genus and probable species (score 2.000–2.299); in yellow: probable genus (score 1.700–1.999), and in red: unreliable results (score ≤1.699). Number of isolates (n) according to identification results are shown in the table below the graph.

4. Conclusions

The Drinking Water Library presented here is composed of 319 MALDI-TOF MS profiles of 120 bacteria isolated from different drinking water sources (distribution networks and mineral water) of different origins, mainly Catalonia (North-East Spain), Andorra and the South of France, and a selection of 199 reference strains of environmental or water-related origin from the Spanish Type Culture Collection. The new library, created using Bruker UltrafleXtreme MALDI-TOF MS equipment, contains MSPs for 164 species of 53 genera already included in the Bruker Daltonics database increasing the species coverage, and 67 possible new species belonging to 44 genera not previously included. Strains of 120 bacteria, including representative strains of potential new species, were deposited at the CECT for public accession, following the established protocols. Further studies are in progress to elucidate the taxonomic status of strains that might belong to new taxa.

The new DWL, which extended the original Bruker Daltonics database with 319 new MSPs for 96 genera from water and other related environments, improved the resolution power for bacterial identification by up to 76%.

Overall, bacterial identification by MALDI-TOF MS is an easy and fast high-throughput technique with low running costs and remarkable specificity. In the present study, there were slight differences in the results between spectrometers, but the method still proved a good choice for interlaboratory collaboration. Moreover, using the formic acid extended direct transfer protocol allowed the operation time to be significantly reduced and when combined with the DWL, it provided a rapid and successful identification of drinking water bacteria. Able to improve identification, MALDI-TOF MS and the DWL constitute an excellent tool for multiple applications, ranging from basic research on water bacterial communities to routine analysis, i.e., in water testing laboratories and water supply companies.

Supplementary material

Additional file 1: tables and figures

CRediT authorship contribution statement

Anna Pinar-Méndez: Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Sonia Fernández:** Conceptualization, Supervision, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **David Baquero:** Conceptualization, Supervision, Formal analysis, Methodology, Writing – review & editing. **Carles Vilaró:** Formal analysis, Methodology, Writing – review & editing. **Belén Galofré:** Conceptualization, Supervision, Writing – review & editing. **Susana González:** Conceptualization, Supervision, Writing – review & editing. **Lidia Rodrigo-Torres:** Formal analysis, Methodology, Writing – review & editing. **David R. Arahal:**

Conceptualization, Supervision, Writing – review & editing. **M. Carmen Macián:** Conceptualization, Supervision, Formal analysis, Methodology, Writing – review & editing. **María A. Ruvira:** Conceptualization, Supervision, Formal analysis, Methodology, Writing – review & editing. **Rosa Aznar:** Conceptualization, Supervision, Writing – review & editing. **Laia Caudet-Segarra:** Formal analysis, Methodology, Writing – review & editing. **Laura Sala-Comorera:** Formal analysis, Methodology, Writing – review & editing. **Francisco Lucena:** Conceptualization, Supervision, Writing – review & editing. **Anicet R. Blanch:** Conceptualization, Supervision, Writing – review & editing. **Cristina Garcia-Aljaro:** Conceptualization, Formal analysis, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was supported by the Spanish Ministry of Science and Innovation (grant number RTC-2015–4496–2) and Pla de Doctorats Industrials of the Catalan Government (2016 DI 083). We thank Dr. Francesc Márquez Garrido (Bruker Daltonics Spain) for technical assistance in constructing the customized Drinking Water Library with MALDI-TOF MS during the three years of the study. We also thank all the technicians and students from the institutions involved in this work for their support: Maria José Arnedo, Ivan Manero, Eva Duran, Daniel Toribio, Judit Martínez, Miquel Jorba, Miriam Navarro, Beatriz Pinto and Rosa Mª Gimenez.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2021.117543.

References

Alatoom, A.A., Cunningham, S.A., Ihde, S.M., Mandrekar, J., Patel, R., 2011. Comparison of direct colony method versus extraction method for identification of gram-positive cocci by use of Bruker biotyper matrix-assisted laser desorption ionization-time of flight mass spectrometry. J. Clin. Microbiol. 49, 2868–2873. https://doi.org/10.1128/JCM.00506-11.

Alcolea-Medina, A., Fernandez, M.T.C., Montiel, N., García, M.P.L., Sevilla, C.D., North, N., Lirola, M.J.M., Wilks, M., 2019. An improved simple method for the identification of Mycobacteria by MALDI-TOF MS (Matrix-Assisted Laser Desorption-Ionization mass spectrometry). Sci. Rep. 9, 5–10. https://doi.org/10.1038/s41598-019-56604-7.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402. https://doi.org/10.1093/nar/25.17.3389.

Angelakis, E., Million, M., Henry, M., Raoult, D., 2011. Rapid and accurate bacterial identification in probiotics and yoghurts by MALDI-TOF mass spectrometry. J. Food Sci. 76, M568–M572. https://doi.org/10.1111/j.1750-3841.2011.02369.x.

Avanzi, I.R., Gracioso, L.H., Baltazar, M.D.P.G., Karolski, B., Perpetuo, E.A., do Nascimento, C.A.O., 2017. Rapid bacteria identification from environmental mining samples using MALDI-TOF MS analysis. Environ. Sci. Pollut. Res. 24, 3717–3726. https://doi.org/10.1007/s11356-016-8125-8.

Bartram, J., Cotruvo, J., Exner, M., Fricker, C., Glasmacher, A., 2003. Expert Consensus in Heterotrophic Plate Counts and Drinking Water Safety (Eds.). IWA Publishing, London, UK, pp. 1–11.

Boers, S.A., Jansen, R., Hays, J.P., 2019. Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. Eur. J. Clin. Microbiol. Infect. Dis. 38, 1059–1070. https://doi.org/10.1007/s10096-019-03520-3.

Chalupová, J., Raus, M., Sedlářová, M., Šebela, M., 2014. Identification of fungal microorganisms by MALDI-TOF mass spectrometry. Biotechnol. Adv. 32, 230–241. https://doi.org/10.1016/j.biotechadv.2013.11.002.

Cunningham, S.A., Patel, R., 2015. Standard matrix-assisted laser desorption ionization-time of flight mass spectrometry reagents may inactivate potentially hazardous bacteria. J. Clin. Microbiol. 53, 2788–2789. https://doi.org/10.1128/JCM.00957-15.

De Carolis, E., Vella, A., Vaccaro, L., Torelli, R., Posteraro, P., Ricciardi, W., Sanguinetti, M., Posteraro, B., 2014. Development and validation of an in-house database for matrix-assisted laser desorption ionization-time of flight mass spectrometry-based yeast identification using a fast protein extraction procedure. J. Clin. Microbiol. 52, 1453–1458. https://doi.org/10.1128/JCM.03355-13.

Drevinek, M., Dresler, J., Klimentova, J., Pisa, L., Hubalek, M., 2012. Evaluation of sample preparation methods for MALDI-TOF MS identification of highly dangerous bacteria. Lett. Appl. Microbiol. 55, 40–46. https://doi.org/10.1111/j.1472-765X.2012.03255.x.

Eddabra, R., Prévost, G., Scheftel, J.M., 2012. Rapid discrimination of environmental Vibrio by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Microbiol. Res. 167, 226–230. https://doi.org/10.1016/j.micres.2011.09.002.

Ghosh, A.K., Paul, S., Sood, P., Rudramurthy, S.M., Rajbanshi, A., Jillwin, T.J., Chakrabarti, A., 2015. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry for the rapid identification of yeasts causing bloodstream infections. Clin. Microbiol. Infect. 21, 372–378. https://doi.org/10.1016/j.cmi.2014.11.009.

Gunnarsdottir, M.J., Gardarsson, S.M., Figueras, M.J., Puigdomènech, C., Juárez, R., Saucedo, G., Arnedo, M.J., Santos, R., Monteiro, S., Avery, L., Pagaling, E., Allan, R., Abel, C., Eglitis, J., Hambsch, B., Hügler, M., Rajkovic, A., Smigic, N., Udovicki, B., Albrechtsen, H.J., López-Avilés, A., Hunter, P., 2020. Water safety plan enhancements with improved drinking water quality detection techniques. Sci. Total Environ. 698, 134185 https://doi.org/10.1016/j.scitotenv.2019.134185.

Hill, V.R., Kahler, A.M., Jothikumar, N., Johnson, T.B., Hahn, D., Cromeans, T.L., 2007. Multistate evaluation of an ultrafiltration-based procedure for simultaneous recovery of enteric microbes in 100-Liter tap water samples. Appl. Environ. Microbiol. 73, 4218–4225. https://doi.org/10.1128/AEM.02713-06.

Janda, J.M., Abbott, S.L., 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory : pluses, perils , and pitfalls. J. Clin. Microbiol. 45, 2761–2764. https://doi.org/10.1128/JCM.01228-07.

Kim, O.S., Cho, Y.J., Lee, K., Yoon, S.H., Kim, M., Na, H., Park, S.C., Jeon, Y.S., Lee, J.H., Yi, H., Won, S., Chun, J., 2012. Introducing EzTaxon-e: a prokaryotic 16s rRNA gene sequence database with phylotypes that represent uncultured species. Int. J. Syst. Evol. Microbiol. 62, 716–721. https://doi.org/10.1099/ijs.0.038075-0.

Kopcakova, A., Stramova, Z., Kvasnova, S., Godany, A., Perhacova, Z., Pristas, P., 2014. Need for database extension for reliable identification of bacteria from extreme environments using MALDI TOF mass spectrometry. Chem. Pap. 68, 1435–1442. https://doi.org/10.2478/s11696-014-0612-0.

Lasch, P., Nattermann, H., Erhard, M., Stämmler, M., Grunow, R., Bannert, N., Appel, B., Naumann, D., 2008. MALDI-TOF mass spectrometry compatible inactivation method for highly pathogenic microbial cells and spores. Anal. Chem. 80, 2026–2034. https://doi.org/10.1021/ac701822j.

Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.P., Göker, M., 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinform. 14, 60. https://doi.org/10.1186/1471-2105-14-60.

Pavlovic, M., Huber, I., Konrad, R., Busch, U., 2013. Application of MALDI-TOF MS for the Identification of food borne bacteria. Open Microbiol. J. 7, 135–141. https://doi.org/10.2174/1874285801307010135.

Rahi, P., Prakash, O., Shouche, Y.S., 2016. Matrix-assisted laser desorption/ionization time-of-flight mass-spectrometry (MALDI-TOF MS) based microbial identifications: challenges and scopes for microbial ecologists. Front. Microbiol. 7, 1–12. https://doi.org/10.3389/fmicb.2016.01359.

Reasoner, D.J., Geldreich, E.E., 1985. A new medium for the enumeration and subculture of bacteria from potable water. Appl. Environ. Microbiol. 49, 1–7. https://doi.org/10.1128/aem.49.1.1-7.1985.

Ruelle, V., El Moualij, B., Zorzi, W., Ledent, P., De Pauw, E., 2004. Rapid identification of environmental bacterial strains by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. 18, 2013–2019. https://doi.org/10.1002/rcm.1584.

Sala-Comorera, L., Blanch, A.R., Vilaró, C., Galofré, B., García-Aljaro, C., 2017. Heterotrophic monitoring at a drinking water treatment plant by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry after different drinking water treatments. J. Water Health 15, 885–897. https://doi.org/10.2166/wh.2017.090.

Sala-Comorera, L., Caudet-Segarra, L., Galofré, B., Lucena, F., Blanch, A.R., García-Aljaro, C., 2020. Unravelling the composition of tap and mineral water microbiota: divergences between next-generation sequencing techniques and culture-based methods. Int. J. Food Microbiol. 334, 108850 https://doi.org/10.1016/j.ijfoodmicro.2020.108850.

Sala-Comorera, L., Vilaró, C., Galofré, B., Blanch, A.R., García-Aljaro, C., 2016. Use of matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) mass spectrometry for bacterial monitoring in routine analysis at a drinking water treatment plant. Int. J. Hyg. Environ. Health 219, 577–584. https://doi.org/10.1016/J.IJHEH.2016.01.001.

Santos, I., Martin, M., Carlton, D., Amorim, C., Castro, P., Hildenbrand, Z., Schug, K., 2017. MALDI-TOF MS for the identification of cultivable organic-degrading bacteria in contaminated groundwater near unconventional natural gas extraction sites. Microorganisms 5, 47. https://doi.org/10.3390/microorganisms5030047.

Santos, I.C., Hildenbrand, Z.L., Schug, K.A., 2016. Applications of MALDI-TOF MS in environmental microbiology. Analyst 141, 2827–2837. https://doi.org/10.1039/c6an00131a.

Sárvári, K.P., Soki, J., Iván, M., Miszti, C., Latkoczy, K., Melegh, S., Urbán, E., 2018. Maldi-TOF MS versus 16S rRNA sequencing: minor discrepancy between tools in identification of bacteroides isolates. Acta Microbiol. Immunol. Hung. 65, 173–181. https://doi.org/10.1556/030.64.2017.025.

Sasser, M., 1990. Identification of bacteria by gas chromatography of cellular fatty acids. USFCC Newsl. 20, 1–6.

Seuylemezian, A., Aronson, H.S., Tan, J., Lin, M., Schubert, W., Vaishampayan, P., 2018. Development of a custom MALDI-TOF MS database for species-level identification of bacterial isolates collected from spacecraft and associated surfaces. Front. Microbiol. 9, 1–8. https://doi.org/10.3389/fmicb.2018.00780.

Tan, K.E., Ellis, B.C., Lee, R., Stamper, P.D., Zhang, S.X., Carroll, K.C., 2012. Prospective evaluation of a matrix-assisted laser desorption ionization-time of flight mass spectrometry system in a hospital clinical microbiology laboratory for identification of bacteria and yeasts: a bench-by-bench study for assessing the impact on time to identification and cost-effectiveness. J. Clin. Microbiol. 50, 3301–3308. https://doi.org/10.1128/JCM.01405-12.

Theel, E.S., Schmitt, B.H., Hall, L., Cunningham, S.A., Walchak, R.C., Patel, R., Wengenack, N.L., 2012. Formic acid-based direct, on-plate testing of yeast and Corynebacterium species by Bruker Biotyper matrix-assisted laser desorption ionization-time of flight mass spectrometry. J. Clin. Microbiol. 50, 3093–3095. https://doi.org/10.1128/JCM.01045-12.

Timperio, A.M., Gorrasi, S., Zolla, L., Fenice, M., 2017. Evaluation of MALDI-TOF mass spectrometry and MALDI BioTyper in comparison to 16S rDNA sequencing for the identification of bacteria isolated from Arctic sea water. PLoS ONE 12, 1–15. https://doi.org/10.1371/journal.pone.0181860.

Tracz, D.M., Antonation, K.S., Corbett, C.R., 2016. Verification of a matrix-assisted laser desorption ionization-time of flight mass spectrometry method for diagnostic identification of high-consequence bacterial pathogens. J. Clin. Microbiol. 54, 764–767. https://doi.org/10.1128/JCM.02709-15.

Tsuchida, S., Umemura, H., Nakayama, T., 2020. Current status of matrix-assisted laser desorption/ionization–time-of-flight mass spectrometry (MALDI-TOF MS) in clinical diagnostic microbiology. Molecules 25, 4775.

Yoon, S.H., Ha, S.M., Kwon, S., Lim, J., Kim, Y., Seo, H., Chun, J., 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int. J. Syst. Evol. Microbiol. 67, 1613–1617. https://doi.org/10.1099/ijsem.0.001755.

Yu, Z., Peruzy, M.F., Dumolin, C., Joossens, M., Houf, K., 2019. Assessment of food microbiological indicators applied on poultry carcasses by culture combined MALDI-TOF MS identification and 16S rRNA amplicon sequencing. Food Microbiol. 82, 53–61. https://doi.org/10.1016/j.fm.2019.01.018.

Lasch, P., Beyer, W., Nattermann, H., Stämmler, M., Siegbrecht, E., Grunow, R., Naumann, D., 2009. Identification of Bacillus anthracis by using matrix-assisted laser desorption ionization-time of flight mass spectrometry and artificial neural networks. Appl. Environ. Microbiol. 75 (22), 7229–7242. https://doi.org/10.1128/AEM.00857-09.