



Differentiation of *Bacillus pumilus* and *Bacillus safensis* Using MALDI-TOF-MS

Raquel Branquinho¹, Clara Sousa^{1,2}, João Lopes^{3,4}, Manuela E. Pintado⁵, Luísa V. Peixe¹, Hugo Osório^{6,7*}

1 REQUIMTE, Laboratório de Microbiologia, Departamento de Ciências Biológicas, Faculdade de Farmácia, Universidade do Porto, Porto, Portugal, **2** CEB, Centro de Engenharia Biológica, Universidade do Minho, Braga, Portugal, **3** REQUIMTE, Laboratório de Química Aplicada, Departamento de Ciências Químicas, Faculdade de Farmácia, Universidade do Porto, Porto, Portugal, **4** CBOF, Centro de Biotecnologia e Química Fina, Escola Superior de Biotecnologia, Universidade Católica Portuguesa, Porto, Portugal, **5** Departamento de Tecnologia Farmacêutica, Faculdade Farmácia, Universidade de Lisboa, Lisboa, Portugal, **6** IPATIMUP, Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Porto, Portugal, **7** Faculdade de Medicina, Universidade do Porto, Porto, Portugal

Abstract

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) despite being increasingly used as a method for microbial identification, still present limitations in which concerns the differentiation of closely related species. *Bacillus pumilus* and *Bacillus safensis*, are species of biotechnological and pharmaceutical significance, difficult to differentiate by conventional methodologies. In this study, using a well-characterized collection of *B. pumilus* and *B. safensis* isolates, we demonstrated the suitability of MALDI-TOF-MS combined with chemometrics to accurately and rapidly identify them. Moreover, characteristic species-specific ion masses were tentatively assigned, using UniProtKB/Swiss-Prot and UniProtKB/TrEMBL databases and primary literature. Delineation of *B. pumilus* (ions at *m/z* 5271 and 6122) and *B. safensis* (ions at *m/z* 5288, 5568 and 6413) species were supported by a congruent characteristic protein pattern. Moreover, using a chemometric approach, the score plot created by partial least square discriminant analysis (PLSDA) of mass spectra demonstrated the presence of two individualized clusters, each one enclosing isolates belonging to a species-specific spectral group. The generated pool of species-specific proteins comprised mostly ribosomal and SASPs proteins. Therefore, in *B. pumilus* the specific ion at *m/z* 5271 was associated with a small acid-soluble spore protein (SASP O) or with 50S protein L35, whereas in *B. safensis* specific ions at *m/z* 5288 and 5568 were associated with SASP J and P, respectively, and an ion at *m/z* 6413 with 50S protein L32. Thus, the resulting unique protein profile combined with chemometric analysis, proved to be valuable tools for *B. pumilus* and *B. safensis* discrimination, allowing their reliable, reproducible and rapid identification.

Citation: Branquinho R, Sousa C, Lopes J, Pintado ME, Peixe LV, et al. (2014) Differentiation of *Bacillus pumilus* and *Bacillus safensis* Using MALDI-TOF-MS. PLOS ONE 9(10): e110127. doi:10.1371/journal.pone.0110127

Editor: Mickaël Desvaux, INRA Clermont-Ferrand Research Center, France

Received: April 17, 2014; **Accepted:** September 17, 2014; **Published:** October 14, 2014

Copyright: © 2014 Branquinho et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: Dr. Kasthuri Venkateswaran, Dr. Irène Ouoba, Dr. Joseph W. Kloepper, Dr. Cecilie From and Dr. Maria Morea are gratefully acknowledged for providing isolates FO-36bT, SAFN-027, SAFN-037, KL-052, 51-3C and 82-2C; Bs31; SE 49 (AP3) and SE 52 (AP7); FEL 55, UNG22 and MIL46, respectively. Raquel Branquinho was supported by a PhD fellowship (Ref. SFRH/BD/61410/2009) and Clara Sousa by a post-doctoral fellowship (Ref. SFRH/BPD/70548/2010), from FCT (Fundação para a Ciência e Tecnologia, Portugal). Hugo Osório acknowledges the funding from QREN-FEDER through the Operational Program ON. 2 – O Novo Norte. IPATIMUP is an Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education and is partially supported by the Portuguese Foundation for Science and Technology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* Email: hosorio@ipatimup.pt

Introduction

Bacillus pumilus and *Bacillus safensis* represent one of the most significant and widespread terrestrial species within the *Bacillus pumilus* group [1,2,3,4,5]. A wide range of biotechnological and pharmaceutical applications has been attributed to these species, including human and animal probiotics [6] or phytosanitary-based products [7]. Moreover, they are important contaminant agents found in industrial settings, namely in food and/or pharmaceutical facilities, posing a serious problem to the quality assurance of these industrial segments [2,4]. More rarely, *B. pumilus* isolates have also been involved in foodborne poisoning [8,9] and in human infections including anthrax-like cutaneous lesions [10,11,12,13].

Studying a collection of previously identified *B. pumilus* isolates we have realized the difficulty to their differentiation from closely related species based on phenotypic and biochemical character-

istics and on 16S rRNA gene sequences [4]. Sequencing housekeeping genes, such as *gyrB* (β -subunit of DNA gyrase) and *rpoB* (β -subunit of RNA polymerase) has proven to be useful for taxonomic resolution of closely related species, including *Bacillus* species. Nevertheless, are already recognized its implementation difficulties in the routine of microbiology laboratories [4,5].

Matrix-assisted laser desorption-ionization-time-of-flight mass spectrometry (MALDI-TOF-MS), is an accurate, fast, and affordable emerging technique that has been increasingly used for identification of bacteria at different taxonomical levels [14,15]. As a proteomic approach, MALDI-TOF-MS relies on the reproducible detection of microbial protein patterns, which can be used for microbial identification by comparing experimental mass spectra with a library of known reference strains or by comparing identities of species-specific biomarkers [16,17].

Using vegetative *B. pumilus* and *B. safensis* cells Farfour et al. [18] unsuccessfully attempted to discriminate these closely related species. Differentiation of these species using spore cells was proposed by Dickinson et al. [19], by MALDI-TOF-MS using a single ion at m/z 7620. Nevertheless, there is a need for further studies in order to assess the validity of this discriminatory tool. Moreover, Lash et al. [20,21] suggested that bacterial vegetative cells when submitted to a protein enrichment process, could provide a more informative mass spectra pattern than that obtained with bacterial spores [20,21].

In addition, the species-specific protein fingerprint constitutes a valuable approach for the assignment of species biomarkers, which usually comprise cell structure and housekeeping proteins [22]. In *Bacillus* spp., ribosomal proteins and small acid-soluble spore proteins (SASPs) have been reported as potential species-specific biomarkers [21]. Nevertheless and despite this successful characterization of *Bacillus cereus* group species [21], this was not applied to the members of the *B. pumilus* group.

In this work, through the implementation of a protocol that enhances proteins extraction [21] combined with chemometric tools, we assessed the potential of MALDI-TOF-MS fingerprinting to discriminate a comprehensive collection of *B. safensis* and *B. pumilus* isolates. Moreover, the tentative assignment of their species-specific protein biomarkers was also performed.

Materials and Methods

Isolates collection and identification

Five *B. pumilus* and twenty-two *B. safensis* isolates that were previously identified by phenotypic and genotypic (16S rRNA, *gyrB* and *rpoB* gene sequences) methods were studied [5,23]. They comprised diverse pulsed-field gel electrophoresis (PFGE)-types and were recovered from different geographic terrestrial locations and sources, including food samples (Norway, Italy and Africa) ($n = 4$), plants (USA) ($n = 2$), gastropods (Portugal) ($n = 3$), health ($n = 6$) and cosmetic ($n = 4$) products (Portugal), and clean room environments from Mars Odyssey (USA) ($n = 5$). Additionally, type and reference strains, *B. safensis* FO-36b^T, *B. pumilus* ATCC 7061^T and ATCC 14884 were included (Table 1).

Sample preparation

For MALDI-TOF-MS analysis, soluble proteins were extracted from a pure single colony of *Bacillus* spp. grown on LB agar (Merck, Darmstadt, Germany), which were subsequently cultured on the same medium, under aerobic conditions, for 24 h at 37°C. Bacterial cells were harvested by transferring three full loops (ca. 30 μ l) from each agar plate into 20 μ l of sterile water and were resuspended by vortexing. Bacterial inactivation was carried out applying the modified trifluoroacetic acid (TFA) (Sigma-Aldrich, St. Louis, MO) inactivation protocol [20], with some modifications previously established by Lash et al. (2008) [20] to improve the accuracy of the mass spectra. Briefly, 80 μ l of pure TFA was added to 20 μ l of each bacterial suspension. After gentle shaking (100 rpm) for 5 min at room temperature, the solution was centrifuged for 20 min at 28960 g^{-1} at 4°C. Subsequently, the supernatant was 10-fold diluted with HPLC grade water (Millipore Corp., Bedford, MA), filtered throughout a 0.22 μ m pore size filter (Millipore) and stored at -20°C until further analysis.

Mass spectrometry methods

Snapshots of different protein composition were detected and acquired by a MALDI-TOF/TOF mass spectrometer (4800 Plus MALDI TOF/TOF Analyzer, AB SCIEX, Framingham, MA), equipped with a 200-Hz frequency Nd:YAG laser, operating at a

wavelength of 355 nm. Pulse ion extraction with a 1300 ns delay time was used for collecting spectra. Measurements were carried out in linear positive mode using an acceleration voltage of 19.4 kV (Grid 1), and a lens 1 voltage of 8 kV. Each spectrum was the accumulated sum of at least 2000 laser shots within the ion range at m/z 2000–12000, due to the good reproducibility of the spectral profile in this interval. All the spectra were externally calibrated using a commercial mixture of angiotensin I, ACTH (adrenocorticotropic hormone) and insulin (AB SCIEX, Framingham, MA) and analyzed with the Data Explorer software (Version 4.6, AB SCIEX, Framingham, MA).

For MALDI-TOF-MS experiments, 2 μ l of the filtrated microbial dilution were mixed with 2 μ l of a 12-mg/ml α -cyano-4-hydroxycinnamic acid (CHCA) (Sigma-Aldrich, St. Louis, MO) solution, prepared in 100% ACN (Acetonitrile, Sigma-Aldrich) and 0.3% TFA. 1 μ l of the mixture was spotted onto a 123×81 mm stainless steel MALDI sample plate (Opti-TOF 384- Well insert, AB SCIEX, Framingham, MA) and allowed to dry at room temperature. For each isolate, two biological replicates (obtained from two different agar plates) were carried out, and the mean spectra were considered for the analysis. Mass spectra were analyzed with the Data Explorer software (v3.7, build 126, AB SCIEX, Framingham, MA). Ion masses were extracted from the raw experimental mass spectra that included all the ion peaks with a relative signal to noise (S/N) ratio intensity above 2.

Chemometric methods

MALDI-TOF-MS spectra were mean-centred and analysed by partial least squares discriminant analysis (PLSDA) [24]. The PLSDA model were developed and validated based on a cross-validation strategy leave-one-out [25] where 70% of the strains were randomly selected to calibrate the model and 30% to test the model (the procedure was repeated 100 times). The PLSDA scores were the source for hierarchical cluster analysis (HCA). The purpose of HCA was the generation of dendrograms highlighting the association between isolates. Dendrograms were performed directly on unprocessed PLSDA scores using the Euclidean distance and the Median's algorithm [26]. All chemometric models were performed in Matlab version 6.5 release 13 (MathWorks, Natick, MA) and the PLS Toolbox version 3.5 for Matlab (Eigenvector Research, Manson, WA).

Biomarker identification

Intact protein masses derived from MS analysis were used to generate a pool of candidate's proteins for the identification of specific markers. The selected distinct mass information was submitted to a web-based TagIdent software tool (<http://web.expasy.org/tagident/>) using 1% mass error for the taxonomic selections *B. pumilus*, *B. safensis* and *B. subtilis*. No restrictions on protein isoelectric point were used. This tool allowed the identification of proteins based on the experimental masses acquired by mass spectrometry using the information available at the UniProtKB/Swiss-Prot and UniProtKB/TrEMBL protein sequence databases. Moreover, ribosomal proteins of genome-sequenced type strains, including *B. pumilus* ATCC 7061^T available in the database developed by Hotta et al. [27] were also included for comparison. For ion peaks matching the theoretical molecular weights we took into consideration the "N-end rule" where N-terminal methionine is cleaved when the second amino acid residue had a small side-chain, according to the previous study [28].

Table 1. Origins of *Bacillus* spp. isolates (n = 27) included in this study.

Isolate	Origin	Year/Location	References
<i>Bacillus pumilus</i>			
Bp ATCC14884			Reference strain
Bp ATCC 7061 ^T			Type Strain
Bp7		2005/Portugal ¹	
Bp11	Health products (n = 3)	2005/Portugal ¹	[37]
Bp15		2005/Portugal ¹	
<i>Bacillus safensis</i>			
Bs1		2004/Portugal ¹	
Bs2	Animals Gastropods (n = 3)	2005/Portugal ¹	
Bs3		2007/Portugal ¹	
Bs13		2005/Portugal ¹	
Bs16	Health products (n = 3)	2005/Portugal ¹	
Bs17		2005/Portugal ¹	[37]
Bs5		2002/Portugal ¹	
Bs18	Cosmetic products (n = 4)	2002/Portugal ¹	
Bs19		2002/Portugal ¹	
Bs27		2002/Portugal ¹	
Bs24		2004/Italy ²	
Bs25	Foods/salame (n = 3)	2004/Italy ²	[38]
Bs33		2004/Italy ²	
Bs22	Plant Growth-Promoting Rhizobacteria (PGPR) (n = 2)	1997/USA ³	[39]
Bs23		1997/USA ³	
Bs31	Food/beans (n = 1)	2003/Africa ⁴	[40]
Bs FO-36b ^T	Clean-room/air particulate (n = 1)	1999/USA ⁵	Type Strain, [2]
Bs35	Clean-room/floor (n = 1)	2001/USA ⁵	
Bs36	Clean-room/cabinet top (n = 1)	2001/USA ⁵	
Bs37	Clean-room/Mars Odyssey spacecraft surface (n = 2)	2001/USA ⁵	[2]
Bs38			
Bs42	Clean-room/anteroom (n = 1)	2001/USA ⁵	

^TType Strain.

¹Isolates obtained from the Quality Control Department (INFARMED), Lisbon, Portugal.

²Isolates FEL 55 from salame felino, UNG22 from salame ungherese and MIL46 from salame milano obtained from the Istituto di Scienze delle Produzioni Alimentari (ISPA), Bari, Italy.

³Isolates SE 49 (AP3) and SE 52 (AP7) from cucumber roots obtained from the Culture collection of the Department of Entomology and Plant Pathology, Auburn University, Alabama, USA.

⁴Isolates Bs31 from African locust beans for *Soumbala* production obtained from Ouagadougou, Africa.

⁵Isolates FO-36b, SAFN-027, SAFN-037, KL-052, 51-3C and 82-2C from spacecraft and assembly-facility surfaces obtained from California Institute of Technology, California, USA.

doi:10.1371/journal.pone.0110127.t001

Results and Discussion

One of the current challenges in bacterial taxonomy is to integrate timely and accurate typing methods for a meaningful identification of microorganisms [14], which is particularly problematic for closely related species such as *B. pumilus* and *B. safensis* [2,5]. Moreover, because of the medical, industrial and biotechnological relevance of these species, reliable, easy and rapid methodologies for their correct differentiation are needed. Discrimination of *B. pumilus* and *B. safensis* by molecular markers sequencing (e.g. *gyrB*), a time demanding and expensive methodology, which is not readily available for routine laboratories, was recently demonstrated [4,5]. In fact, there are no comprehensive studies assessing its potential on *B. pumilus* and *B.*

safensis discrimination despite the importance of the MALDI-TOF-MS application in bacterial differentiation. Moreover, combining accurate MALDI-TOF mass ion signals with the information available at a protein sequence database, such as UniProt, species-specific candidate protein biomarkers can be tentatively assigned, supporting the interest of this methodology for species identification.

Sample preparation conditions

Bacterial species identification using MALDI-TOF-MS is based on mass profiles obtained from whole bacterial cell suspensions considering proteins with low mass weight (less than 20 kDa). Additionally, for MS applications it is imperative to define a standardized protocol, including the establishment of rigorous

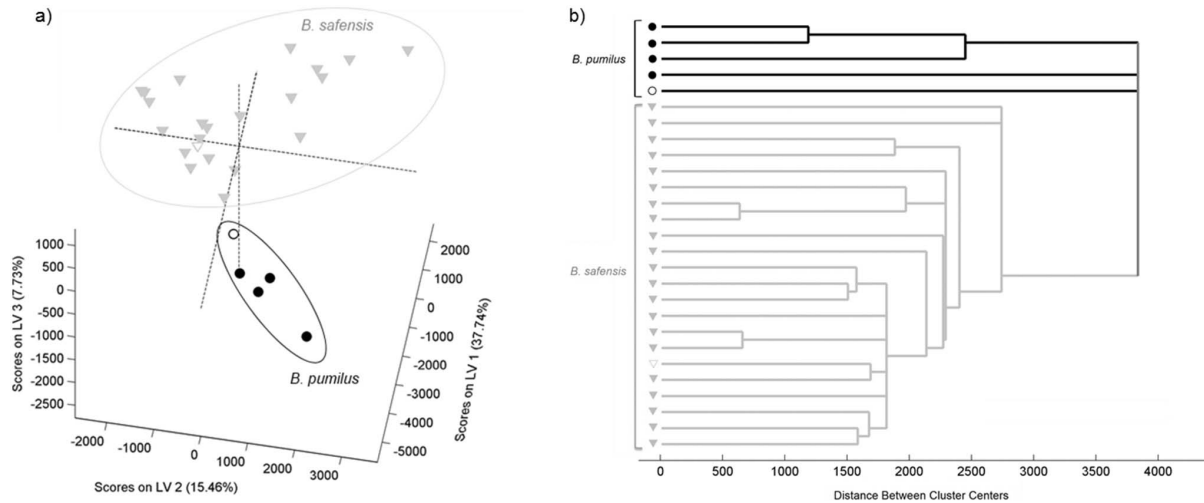


Figure 1. Average mass spectra obtained by MALDI-TOF-MS analysis of *B. pumilus* and *B. safensis* isolates. (a) *B. pumilus* (5 spectra) and (b) *B. safensis* (22 spectra) in the ion range at *m/z* 2000 to 12000. Spectra were obtained by averaging the respective experimental ion signals from all isolates.
doi:10.1371/journal.pone.0110127.g001

sample preparation and cultivation conditions, including culture medium, temperature and time. Therefore, an adequate number of ion masses should be obtained to allow the identification and discrimination of closely related bacterial species. Moreover, the definition of strict parameters for spectral data acquisition is also required.

The protein enrichment protocol, reported by Lash et al. [20], for *B. cereus* group members, using a combined TFA treatment, centrifugation and filtration steps, is of relevance in the case of *Bacillus* species, since it promoted an efficient extraction of soluble microbial proteins presented in the core spore and other morphological structures [20]. The described sample preparation procedure was successfully applied to *B. pumilus* and *B. safensis*

isolates, clonally diverse and collected from different terrestrial origins [4], demonstrating its suitability to generate reproducible mass spectra data with a sufficient number of ion masses and reinforcing the potential for the successful application of the MS technique for their identification in routine laboratories.

Mass spectrometry analysis

The Figure 1 displayed the average mass spectra obtained from MALDI-TOF-MS analysis for the two *Bacillus* species under study: (a) *B. pumilus* (5 spectra) and (b) *B. safensis* (22 spectra), in the ion range at *m/z* 2000 to 12000. In fact, MALDI-TOF-MS profile analysis revealed the presence of species-specific ion signals.

Table 2. Species-specific ion *m/z* values (average) of *B. pumilus* and *B. safensis* isolates.

Experimental average ion <i>m/z</i> values ^a		
	<i>B. pumilus</i>	<i>B. safensis</i>
		3692.5
<i>B. pumilus</i>	3060 (BP1) ^b	3821.5
	3608.5 (BP2)	4305.5
	5271 (BP3)	5948.5
	6122 (BP4)	6704
		6793.5
		7415
	3692.5	
	3821.5	3396 (BS1) ^b
	4305.5	5288 (BS2)
<i>B. safensis</i>	5948.5	5568 (BS3)
	6704	6094 (BS4)
	6793.5	6413 (BS5)
	7415	

^a average ion *m/z* values (± 2 Da).
^b In brackets were named candidate species-specific ion masses.
doi:10.1371/journal.pone.0110127.t002

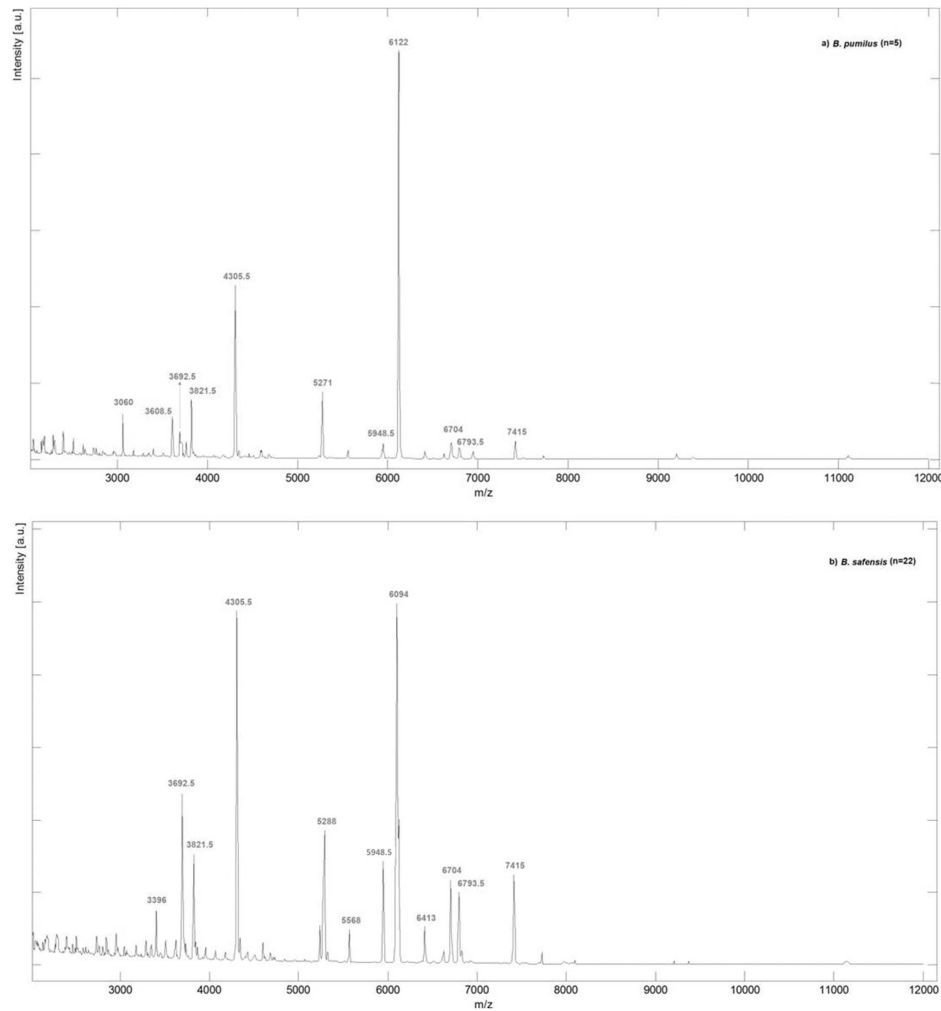


Figure 2. Score plot obtained by PLSDA regression model and the corresponding dendrogram. Score plot of the PLSDA regression model (a) and respective dendrogram (b), of *B. pumilus* and *B. safensis* isolates. Legend: ● *B. pumilus* isolates and ▼ *B. safensis* isolates. Unfilled symbols correspond to the type strains of both species. doi:10.1371/journal.pone.0110127.g002

Table 2 compiled the characteristic ion masses for each spectral group and those presented in both species.

As expected, similar fingerprints could be observed for isolates belonging to the same species, exhibiting several common ion peaks (Figure 1). For instance, the ions having *m/z* values at 3060, 3608.5, 5271 and 6122 were consistently observed in *B. pumilus* isolates, while the ions having the *m/z* values at 3396, 5288, 5568, 6094 and 6413 were related to *B. safensis*. Moreover, *B. pumilus* and *B. safensis* isolates demonstrated common ions at *m/z* 3692.5, 3821.5, 4305.5, 5948.5, 6704, 6793.5 and 7415 (Table 2), which could be considered characteristic ion peaks for both species. Therefore, the protein fingerprint similarities achieved among *B. pumilus* and *B. safensis* corroborated the closeness similarity previously verified among them [4,5].

Interestingly, some spectral variability was observed within isolates belonging to the same spectral group, including the presence or absence of some ion peaks beyond those listed in table 2 (data not shown). This observation was not surprising, since we had previously demonstrated that both species comprised a clonally diverse population [4]. Therefore, the specific peptide profile probably reflects their evolution towards an adaptation to different niches [4].

Indeed, few studies have explored the potential of MALDI-TOF-MS to properly identify *B. pumilus* or *B. safensis* isolates [18,19,29,30].

Farfour et al. attempted unsuccessfully to discriminate between *B. pumilus* and *B. safensis* vegetative cells by MALDI-TOF-MS using the Andromas database [18]. This discrepancy highlighted the need for the improvement and enlargement of this database and led eventually to the prior enrichment of the protein before MS analysis. Moreover, Böhme et al. [29], using *B. pumilus* type and reference strains (ATCC 7061^T and 14884) that were subjected to an extraction procedure previously to the MS analysis, suggested the presence of a series of ions at *m/z* 3620, 5297, 6617 and 7237, which were specific for this species, when compared with other *B. subtilis* group members (*B. subtilis*, *Bacillus amyloliquefaciens* and *Bacillus licheniformis*) and with members of *Bacillus cereus* group (*B. cereus*, *Bacillus megaterium* and *Bacillus thuringiensis*). In addition, the presence of this same series of ions at *m/z* 3620, 6617 and 7238 were also detected by Fernández-No et al. [30] in the same reference strains. Analysis of our MALDI-TOF-MS profiles did not reveal these ion peaks as *B. pumilus* species-specific discriminatory, when compared with *B. safensis*. Nevertheless, a closer inspection of the mass spectra of *B.*

Table 3. Overview of biomarkers tentative assignment of MALDI-TOF-MS mass signals of *B. pumilus* group species.

Specie(s)	Observed Mass (Da)	Predicted Mass (Da)	Error (%)	UniProt Accession ID	Protein Description	Peptide sequence	Organism ^e
	3060	NA	NA	Unassigned	NA		NA
	3608.5	NA	NA	Unassigned	NA		NA
	5271	5270	0.02	A8FJG4	50S ^h RP subunit L34	MKRTFQNNRKRKSVHGFRRMSSKNGRLVLRKRBSKGRKKLSA	<i>B. pumilus</i> SAFR-032
<i>B. pumilus</i>		5266.89	0.08	A8FDQ9	^a SASP O	MTKRKANHVINGMNAAKSQNGAGYIEDDQLVLTAEQRQNNKRKKNQ	<i>B. pumilus</i> SAFR-032
		6114	0.2	C0H3U0	Uncharacterized membrane protein YyzG	MQTRVILLAVMICLVSAITVYLLNGCKYDFLDIGGTIGCGLGFVVRVQKKQS	<i>B. subtilis</i> subsp. <i>subtilis</i> str. 168
	6122	6126	0.08	P0C8M5	Transcriptional regulator SfrA	MKTHVKKDLDKGWHMLIQEARSIGLGIHDVROFLESETASRKKNHKKTVRQD	<i>B. subtilis</i> subsp. <i>subtilis</i> str. 168
	3049	NA	NA	Unassigned	NA		NA
	3396	NA	NA	Unassigned	NA		NA
<i>B. safensis</i>		5288	5299.88*	A8FHD4	^a SASP J	MSFFQDKKAKSEKDKHQVDQLLEEAASKELAGDPLQEAQVQKKKNDQ	<i>B. pumilus</i> SAFR-032
		5568	5544	A8FDQ8	^a SASP P	MTNKNKTKDIRQNSPKHEHSGOPEPLSGSKYKRNHTRQKHNSHHDM	<i>B. pumilus</i> SAFR-032
	6094	NA	NA	Unassigned	NA		NA
	6413	6411.7*	0.03	A8FCW7	^h RP subunit L32	MAVPRRTSKMKKRLRRTHFKLOVPGMVAACPECEGEMKISHRVCKSCGYKGDWKS	<i>B. pumilus</i> ATCC 7061 ^T
	3692.5	3698	0.1	C0H3V1	UPF0752 membrane protein YczN	MSGYSNGGGYGGISSFALLIWLIFILLIYGTAFVGGF	<i>B. subtilis</i> subsp. <i>subtilis</i> str. 168
	3821.5	NA	NA	Unassigned	NA		NA
	4305.5	4305	0.01	A8F9A9	50S ^h RP subunit L36	MKVRPSVKPICEKCKVIRRKGVMMVICENPKHKQKQG	<i>B. pumilus</i> SAFR-032
<i>B. pumilus</i> and <i>B. safensis</i>		5948.5	5932	A8FF72	50S ^h RP subunit L33 2	MRVNIILACTECCGERNYITKKNRNNPDRVEFKKYSRDKKQTVHRETK	<i>B. pumilus</i> SAFR-032
	6704	6691	0.3	C0H3Z1	Uncharacterized protein YyzG	MMKNGFAYKNGKLVNIFCGKEELYNELKAFVYKTSINNVKEVSRPSYRRTKSKQLE	<i>B. subtilis</i> subsp. <i>subtilis</i> str. 168
	6793.5	6793*	0.01	B4AE40	50S ^h RP subunit L28	MARKCVITGRKTKAGNNRSHAMNSTRKTWGANLQKVRILVDGPKPR VYVSARALKSGKVERV	<i>B. pumilus</i> ATCC 7061 ^T
	7415	7410.8*	0.08	B4AM90	50S ^h RP subunit L35	MPKMKTHRGSARFKKTKGSKLKRSHAYTSHLF ANKSTKQKRKLRKSAIVSAGDFKRIKQQLANIK	<i>B. pumilus</i> ATCC 7061 ^T

Protein identity was determined by the TagIdent software and compared with ribosomal subunit proteins developed by Hotta et al. [23] described in Materials and Methods section.

^aSASP - Small, acid-soluble spore protein.

^hRP - Ribosomal protein.

^eOrganism - bacterial strain where the protein was described.

NA - not applicable.

*Predicted molecular weight (Mw) proposed by Hotta et al. [27], considering "N-end rule" [28] where N-terminal methionine is cleaved from specific penultimate amino acid residues such as glycine, alanine, serine, proline, valine, threonine and cysteine.

doi:10.1371/journal.pone.0096944.t003

Table 4. Candidate species-specific biomarkers assignments of *B. pumilus* and *B. safensis*.

Ion mass (<i>m/z</i>)	Species	
	<i>B. pumilus</i>	<i>B. safensis</i>
5271	+	–
5288	–	+
5568	–	+
6122	+	–
6413	–	+

Ion masses are presented as *m/z* values. The presence/absence of an ion peak in each spectral group is represented by +/-, respectively.
doi:10.1371/journal.pone.0110127.t004

pumilus ATCC 7061^T and 14884, revealed the presence of ions having *m/z* values at 3621, 5290 and 6624, although the ion at *m/z* 7238 was not present. Indeed, the differences found in these *B. pumilus* profiles could be justified by the distinct growth culture medium used and the sample preparation procedure employed.

Finally, in other work, Dickinson et al. [19] presented MALDI-TOF-MS as a useful taxonomic tool for differentiating spores of *B. pumilus* and *B. safensis*. Results revealed the presence of two groups of characteristic ion peaks, comprising *B. pumilus* (ions at *m/z* 6860, 7230 and 9606) and *B. safensis* (ions at *m/z* 6860, 7230, 7620 and 9606). The authors claimed the presence of the additional ion at *m/z* 7620 in the spectra profile of *B. safensis* as a species-specific biomarker, allowing the discrimination of these two species. Nonetheless, analysis of our *B. safensis* spectral data (n = 22) did not reveal the presence of this ion peak, probably because the obtained MALDI-TOF-MS profiles were from vegetative cells. Additionally, the spectral profile obtained from spores seems to be more laborious and insufficient to discriminate appropriately among these closely related species, since spectral data with few number of ion peaks were generated. For these reasons, it was not possible to compare our study with the existing ones as we have used different cultural conditions and sample preparation procedures.

On the other hand, the species identification is limited to the bacterial species spectrum presented in a specific MS database. Therefore, bacterial identification is only possible inside the frame of bacterial reference spectrum of the database used. Indeed, few well-characterized *B. pumilus* isolates are available in public databases, as the SpectraBank (<http://www.spectrabank.org>), namely the type strain ATCC 7061^T and the reference strain ATCC 14884, and no *B. safensis* was yet included, which constrains its identification.

Chemometric analysis

B. pumilus and *B. safensis* were clearly discriminated by a combined MALDI-TOF-MS and chemometric approach. The score plot generated by PLS-DA of mass spectra of all isolates tested exhibited two individualized clusters, each one enclosing isolates belonging to a particular spectral group, which included reference and type strains (Fig. 2) of (a) *B. pumilus* and (b) *B. safensis*. The PLS-DA scores were also presented as a dendrogram corroborating the species discrimination into two distinct clusters (Fig. 2). Moreover, this approach allowed the discrimination of the two *Bacillus* species with 100% of sensitivity and specificity.

Despite the recognized proficiency of chemometric tools for the analysis and identification of bacteria based on their fingerprint, in the case of *Bacillus* spp., this characterization was only previously applied in *B. cereus* group species [21]. Therefore, this is the first

successful application of this approach considering members of *B. pumilus* group, stressing its relevance for discrimination among close related species.

Candidate molecular biomarkers assignment

The possibility of biomarkers identification is one of the most valuable aspects of the mass spectrometric-based identification techniques, being this approach successfully applied to different bacterial species [21,31,32].

The assignment of 17 mass signals diagnostic ions formed in the MALDI-TOF-MS of *B. pumilus* and *B. safensis* represented the first consistent evidence of the relation between these ion *m/z* signals and the specific candidate protein sequences, which were presented in Table 3. Direct bacterial discrimination by means of MALDI-TOF-MS was hampered by the absence of consistent databases supported on sufficient identified biomarkers. Indeed, only two *B. pumilus* genomes with numerous proteins defined as unknown were available in UniProtKB/Swiss-Prot and UniProtKB/TrEMBL (*B. pumilus* SAFR-032 and ATCC 7061^T) and no *B. safensis* were deposited, which hindered the candidate biomarkers identification. Therefore, this type of assignments can only be tentatively used to establish potential connections between protein sequences and ion *m/z* signals, and thus, should be prudently interpreted, although previously successfully applied in *B. cereus* group members [21].

We found evidences of *B. pumilus* and *B. safensis* specific biomarkers, associated with a series of ions at *m/z* 4305.5, 5948.5, 6793.5 and 7415, which was attributed to the 50S ribosomal subunits proteins, respectively, L36, L33, L28 and L35 of *B. pumilus* SAFR-032 (correspondent amino acidic sequences were also presented in Table 3). Moreover, the remaining ion detected at *m/z* of 3821.5 was not assigned, and two showed correspondences with membrane proteins of *B. subtilis* subsp. *subtilis* str 168, the YczN and YjzG at *m/z* of 3692.5 and 6704, respectively.

The tentative assignment of *B. pumilus* specific biomarkers revealed the possible correspondence of the diagnostic ion at *m/z* 5271 with the 50S ribosomal subunit protein L34 or with the SASP O. In addition, the characteristic ion at *m/z* 6122 was diagnostic for either the uncharacterized membrane protein (YjzG) and/or the transcriptional regulator – SlrA of *B. subtilis* subsp. *subtilis* str 168. Additionally, the characteristic ions at *m/z* 3060 and 3608.5 were not possible to assign.

Concerning *B. safensis* specific ions at *m/z* 5288, 5568 and 6413, potentially corresponding with two specific SASPs (SASP J and SASP P) and a 50S ribosomal subunit protein L32 were also found in *B. pumilus* SAFR-032. The remaining ion *m/z* peaks detected were not possible to designate. Therefore, the proposed characteristic biomarkers, which could be used to differentiate

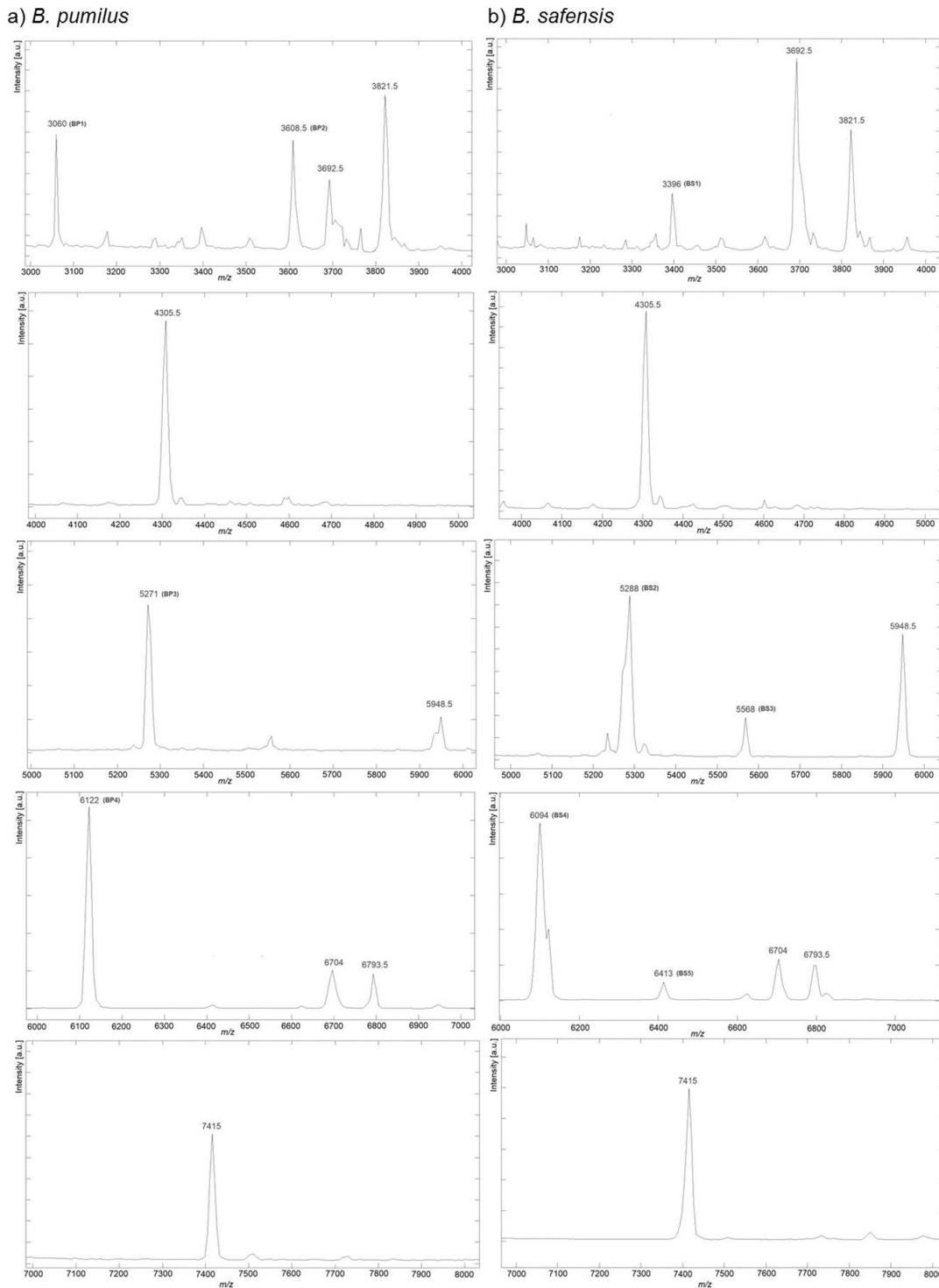


Figure 3. MS patterns of the candidate species-specific ion peaks (m/z) using MALDI-TOF-MS in linear mode. a) *B. pumilus* (average of 5 spectra) and b) *B. safensis* (average of 22 spectra). In brackets were named candidate species-specific ion peaks. doi:10.1371/journal.pone.0110127.g003

between *B. pumilus* and *B. safensis* are summarized in Table 4. Moreover, these differentiating series of ions were shown in Figure 3, which outlined the representative MALDI-TOF-MS biomarkers established for *B. pumilus* and *B. safensis*.

In fact, ribosomal proteins and small, acid-soluble spore proteins (SASPs) have been suggested to be responsible for many

ion masses detected by MALDI-TOF-MS profiles [21]. Since up to 21% of the overall cellular protein content is ribosomal and because ribosomal proteins are part of the cellular translational machinery constitutively expressed in vegetative cells, they constitute a stable ensemble of protein biomarkers suitable for use by fingerprinting techniques [33]. Moreover SASPs, a group of

species-specific proteins present in large amounts in the core region of *Bacillus* endospores, have been also proposed as biomarkers for rapid differentiation and identification of *Bacillus* spp. using mass spectrometry approaches [34,35,36].

Our results also suggested that ribosomal and spore proteins constituted most of the *B. pumilus* and *B. safensis* biomarkers. A more detailed analysis could be carried out with MS/MS peptide fragmentation of the specific proteins assigned and subsequent comparison in protein databases or even with MS/MS peptide *de novo* sequencing. Nevertheless, within the context of the present work, which aimed to establish a MALDI-TOF-MS fingerprint classification for *B. pumilus* and *B. safensis*, these results may be beneficial and improve further accuracy of MS-based detection methods in identifying these species.

Conclusion

MALDI-TOF-MS profiles combined with chemometric analysis (PLSDA) proved to be valuable tools for discrimination of *B.*

pumilus and *B. safensis*, allowing its rapid identification. These high throughput approaches should be promptly considered for *Bacillus* species identification due to the inaccuracy of conventional techniques in the identification of closely related species of this genus. In this sense, it is imperative to standardize a sample preparation protocol, which should include a protein extraction and enrichment step, to provide informative and reproducible mass spectra. Moreover, tentative assignment of *B. pumilus* and *B. safensis* protein biomarkers suggested that most of them are ribosomal and spore proteins.

Author Contributions

Conceived and designed the experiments: RB CS JL LP HO. Performed the experiments: RB CS HO. Analyzed the data: RB JL CS MP LP HO. Contributed reagents/materials/analysis tools: CS JL MP LP HO. Contributed to the writing of the manuscript: RB CS LP HO.

References

1. U. S. Environmental Protection Agency (USEPA) (2004) Biopesticides Registration Action Document, *Bacillus pumilus* strain QST 2808 (PC Code 006485). Office of Pesticide Programs, Washington, DC: Government Printing Office.
2. Satomi M, La Duc MT, Venkateswaran K (2006) *Bacillus safensis* sp. nov., isolated from spacecraft and assembly-facility surfaces. *Int J Syst Evol Microbiol* 56: 1735–1740.
3. Pérez-García A, Romero D, de Vicente A (2011) Plant protection and growth stimulation by microorganisms: biotechnological applications of *Bacilli* in agriculture. *Curr Opin Biotechnol* 22: 187–193.
4. Branquinho R, Meirinhos-Soares L, Carricho JA, Pintado M, Peixe LV (2014) Phylogenetic and clonality analysis of *Bacillus pumilus* isolates uncovered a highly heterogeneous population of different closely related species and clones. *FEMS Microb Ecol*. doi:10.1111/1574-6941.12426
5. Liu Y, Lai Q, Dong C, Sun F, Wang L, et al. (2013) Phylogenetic diversity of the *Bacillus pumilus* group and the marine ecotype revealed by Multilocus Sequence Analysis. *PLoS ONE* 8: 1–11.
6. Hong HA, Duc le H, Cutting SM (2005) The use of bacterial spore formers as probiotics. *FEMS Microbiol Rev* 29: 813–835.
7. Pérez-García A, Romero D, de Vicente A (2011) Plant protection and growth stimulation by microorganisms: biotechnological applications of *Bacilli* in agriculture. *Curr Opin Biotechnol* 22: 187–193.
8. From C, Hormazabal V, Granum P (2007) Food poisoning associated with pumilacidin-producing *Bacillus pumilus* in rice. *Int J Food Microbiol* 115: 319–324.
9. From C, Pukall R, Schumann P, Hormazabal V, Granum PE (2005) Toxin-producing ability among *Bacillus* spp. outside the *Bacillus cereus* group. *Appl Environ Microbiol* 71: 1178–1183.
10. Haymore BR, Akers KS, Ferguson TM (2006) A case of persistent *Bacillus pumilus* bacteremia associated with cholangitis. *J Infect* 52: 154–155.
11. Bentur HN, Dalzell AM, Riordan FAI (2007) Central venous catheter infection with *Bacillus pumilus* in an immunocompetent child: a case report. *Ann Clin Microbiol Antimicrob* 6: 12.
12. Tena D, Martínez-Torres JA, Pérez-Pomata MT, Sáez-Nieto JA, Rubio V, et al. (2007) Cutaneous infection due to *Bacillus pumilus*: report of 3 cases. *Clin Infect Dis* 44: 40–42.
13. Johnson BT, Shaw LN, Nelson DC, Mayo JA (2008) Extracellular proteolytic activities expressed by *Bacillus pumilus* isolated from endodontic and periodontal lesions. *J Med Microbiol* 57: 643–651.
14. Clark AE, Kaleta EJ, Arora A, Wolk DM (2013) Matrix-assisted laser desorption ionization-time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. *Clin Microbiol Rev* 3: 547–603.
15. Keys CJ, Dare DJ, Sutton H, Wells G, Lunt M, et al. (2004) Compilation of a MALDI TOF mass spectral database for the rapid screening and characterization of bacteria implicated on human infectious diseases. *Infect Genet Evol* 3: 221–42.
16. Mellmann A, Bimet F, Bizet C, Borovskaya AD, Drake RR, et al. (2009) High interlaboratory reproducibility of matrix-assisted laser desorption ionization-time of flight mass spectrometry-based species identification of nonfermenting bacteria. *J Clin Microbiol* 47: 3732–4.
17. Emonet S, Shah HN, Cherkaoui A, Schrenzel J (2010) Application and use of various mass spectrometry methods in clinical microbiology. *Clin Microbiol Infect* 11: 1604–1613.
18. Farfour E, Leto J, Barritault M, Barberis C, Meyer J, et al. (2012) Evaluation of the Andromas matrix-assisted laser desorption ionization-time of flight mass

- spectrometry system for identification of aerobically growing Gram-positive bacilli. *J Clin Microbiol* 50: 2702–7.
19. Dickinson DN, La Duc MT, Satomi M, Winefordner JD, Powell DH, et al. (2004) MALDI-TOF MS compared with other polyphasic taxonomy approaches for the identification and classification of *Bacillus pumilus* spores. *J Microbiol Methods* 58: 1–12.
20. Lasch P, Nattermann H, Erhard M, Stämmler M, Grunow R, et al. (2008) MALDI-TOF mass spectrometry compatible inactivation method for highly pathogenic microbial cells and spores. *Anal Chem* 80: 2026–2034.
21. Lasch P, Beyer W, Nattermann H, Stämmler M, Siegbrecht E, et al. (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: 7229–42.
22. Welker M, Moore ER (2011) Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. *Syst Appl Microbiol* 34: 2–11.
23. Wang LT, Lee FL, Tai CJ, Kasai H (2007) Comparison of *gyrB* gene sequences, 16S rRNA gene sequences and DNA–DNA hybridization in the *Bacillus subtilis* group. *Int J Syst Evol Microbiol* 57: 1846–1850.
24. Barker M, Rayens W (2003) Partial least squares for discrimination. *J Chemom* 17: 166–173.
25. Sousa C, Grosso F, Meirinhos-Soares L, Peixe L, Lopes J (2014) Identification of carbapenem-resistant *Acinetobacter baumannii* clones using infrared spectroscopy. *Journal of Biophotonics* 7: 287–294.
26. Næs T, Isaksson T, Fearn T, Davis T (2002) A user-friendly guide to multivariate calibration and classification. NIR Publications, Chichester UK.
27. Hotta Y, Sato J, Sato H, Hosoda A, Tamura H (2011) Classification of the genus *Bacillus* based on MALDI-TOF MS analysis of ribosomal proteins coded in S10 and *spe* operons. *J Agric Food Chem* 59: 5222–30.
28. Sherman F, Stewart JW, Tsunasawa S (1985) Methionine or not methionine at the beginning of a protein. *Bioessays* 1: 27–31.
29. Böhme K, Fernández-No IC, Barros-Velázquez J, Gallardo JM, Cañas B, et al. (2011) Rapid species identification of seafood spoilage and pathogenic Gram-positive bacteria by MALDI-TOF mass fingerprinting. *Electrophoresis* 32: 2951–65.
30. Fernández-No IC, Böhme K, Díaz-Bao M, Cepeda A, Barros-Velázquez J, et al. (2013) Characterization and profiling of *Bacillus subtilis*, *Bacillus cereus* and *Bacillus licheniformis* by MALDI-TOF mass fingerprinting. *Food Microbiol* 33: 235–242.
31. Dieckmann R, Malorny B (2011) Rapid Screening of Epidemiologically Important *Salmonella enterica* subsp. *enterica* Serovars by Whole-Cell Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry. *Appl Environ Microbiol* 77: 4136–4146.
32. Momo RA, Povey JF, Smales CM, O'Malley CJ, Montague GA, et al. (2013) MALDI-TOF mass spectrometry coupled with multivariate pattern recognition analysis for the rapid biomarker profiling of *Escherichia coli* in different growth phases. *Anal Bioanal Chem* 25: 8251–8265.
33. Arnold RJ, Reilly JP (1999) Observation of *Escherichia coli* ribosomal proteins and their posttranslational modifications by mass spectrometry. *Anal Biochem* 269: 105–112.
34. Driks A (1999) The *Bacillus subtilis* spore coat. *Microbiol Mol Biol Rev* 63: 1–20.
35. Driks A (2002) Proteins of the spore core and coat. p. 527–535. In Sonenshein AL, Hoch JA, Losick R, *Bacillus subtilis* and its Closest Relatives: from Genes to Cells, ASM Press, Washington, DC.

36. Lai EM, Phadke ND, Kachman MT, Giorno R, Vazquez S, et al. (2003) Proteomic analysis of the spore coats of *Bacillus subtilis* and *Bacillus anthracis*. *J Bacteriol* 185: 1443–1454.
37. Branquinho R, Pintado ME, Peixe L (2012) Clonality and protein diversity of *Bacillus pumilus* isolates from different sources and geographic regions. *Planta Med* 78: PA5, doi:10.1055/s-0032-1320320
38. Matarante A, Baruzzi F, Cocconcelli PS, Morea M (2004) Genotyping and toxigenic potential of *Bacillus subtilis* and *Bacillus pumilus* strains occurring in industrial and artisanal cured sausages. *Appl Environ Microbiol* 70: 5168–5176.
39. Jetiyanon K (1997) Interaction between PGPR and cucumber during induced systemic resistance: recognition and early host defense responses. Ph.D. dissertation, Auburn University, Auburn, AL, USA.
40. Ouoba L, Diawara B, Amoa-Awua WK, Traore AS, Moller PL (2004) Genotyping of starter cultures of *Bacillus subtilis* and *Bacillus pumilus* for fermentation of African locust bean (*Parkia biglobosa*) to produce Soubala. *Int J Food Microbiol* 90: 197–205.